

Джеймс Гудвин

ИССЛЕДОВАНИЕ В ПСИХОЛОГИИ

МЕТОДЫ И ПЛАНИРОВАНИЕ

3-е издание



C. James Goodwin

RESEARCH IN PSYCHOLOGY

METHODS AND DESIGN

Third Edition



John Wiley & Sons, Inc.

Джеймс Гудвин

ИССЛЕДОВАНИЕ В ПСИХОЛОГИИ

МЕТОДЫ И ПЛАНИРОВАНИЕ

3-е издание



Москва · Санкт-Петербург · Нижний Новгород · **Воронеж**
Ростов-на-Дону · Екатеринбург · **Самара**
Киев • Харьков • Минск
2004

ББК 88.3
УДК 159.9.072
Г93

Г93 Исследование в психологии: методы и планирование / Дж. Гудвин. — 3-е изд. — СПб.: Питер, 2004. — 558 с: ил. — (Серия «Мастера психологии»).

ISBN 5-94723-290-1

Предлагаемая вниманию читателей книга представляет собой учебник, включающий материалы курса экспериментальной психологии. Книга состоит из 12 глав, эпилога и нескольких полезных приложений. От других книг подобного рода это издание выгодно отличает ряд особенностей, связанных с организацией и подачей материала, облегчающих студентам изучение данной дисциплины. Каждая глава открывается кратким обзором и постановкой учебных задач и завершается обобщением важной информации, проверочными заданиями и упражнениями. Книга предназначена для студентов и преподавателей психологических факультетов, а также для всех интересующихся методами психологического исследования.

ББК 88.3
УДК 159.9.072

ISBN 0-471-39861-6 (англ.)

ISBN 5-94723-290-1

Оглавление

Предисловие

Философия книги

При разработке трех изданий этой книги я руководствовался несколькими идеями. Во-первых, желанием показать студентам ход мысли психологов-экспериментаторов, поэтому читатели найдут здесь подробное описание сущности психологической науки и ее отличий от псевдонауки, логики научного мышления и подходов, применяющихся психологами для а) планирования исследований, б) их проведения и в) разработки обоснованных выводов. Во-вторых, необходимостью показать студентам, что есть многие методы объяснения поведения, поэтому хотя основной темой книги является метод эксперимента, вы найдете в ней всестороннее рассмотрение других исследовательских подходов. В-третьих, будучи уверенным в том, что исследователи не должны забывать об этической стороне своих экспериментов, я поместил в книгу отдельную главу, посвященную этике (глава 2), и включил дополнительные разделы на данную тему в остальные главы книги. В-четвертых, я уделяю большое внимание истории психологии и считаю, что невозможно понять настоящее без знания прошлого, поэтому в книгу вошла информация из истории экспериментальной психологии. Однако понимая, что целью данного курса является изучение современных методов исследования, а не истории психологии, я отобрал исторические данные, освещающие только важные методологические вопросы. В-пятых, и, возможно, именно это является самым главным, я считаю, что проведение психологических исследований — это весьма увлекательное занятие, но заметил, что некоторые студенты не испытывают особого энтузиазма. Они предполагают, что данный курс будет скучным, трудным и не особенно полезным для них. Поэтому я постарался создать книгу, интересную для студентов — привлекательную (со множеством интересных описаний реальных исследований), понятную (написанную доступным и приближенным к разговорному языком) и полезную (способствующую развитию необходимых навыков критического мышления).

Организация материала

Книга состоит из 12 глав, эпилога и нескольких полезных приложений. Глава 1 является основой всего дальнейшего изложения и содержит подробное описание научного мышления в сравнении с ненаучным и псевдонаучным. Глава 2 посвящена этике исследований и вопросам применения этического кодекса Американской психологической ассоциации *(APA)* в исследованиях с испытуемыми людьми или животными, также в этой главе обсуждается научное мошенничество. В главе 3 рассматривается вопрос возникновения идеи исследования и постепенного развития взаимосвязи этого исследования с разработанной ранее теорией. Также даются основы использования наиболее важной психологической электронной базы данных — *PsycINFO.* В главе 4 рассказывается об измерении

поведения, создании выборок и статистическом анализе данных. Далее следуют четыре главы, посвященные экспериментальным методам — введение в экспериментальные методы (глава 5), обсуждение решения проблем экспериментальных исследований (глава 6), а также две главы об экспериментальных планах (главы 7 и 8). В следующих главах описываются различные виды психологических исследований, в том числе корреляционные исследования (глава 9), квазиэкспериментальные планы, прикладные исследования и оценка программы (глава 10), исследования с использованием планов с «небольшим ЛГ» (глава 11) и описательные исследования (глава 12). В приложениях рассказывается о подготовке (не)известного типа отчета *APA,* составлении опросов, проведении статистического анализа и о том, как сделать выводы о результатах исследования. Последнее приложение предоставляет ответы к упражнениям и тестам, помещенным в конце каждой главы.

В изложение включены вставки трех типов: «История», «Классические исследования» и «Этика». Во вставках типа «История» приводится интересная информация об исторических корнях экспериментальной психологии и рассказывается о создании и развитии различных методов исследований. В «Классических исследованиях» описываются широко известные эксперименты (например, эксперимент с куклой бобо), иллюстрирующие конкретные исследовательские планы и методы. Вставки типа «Этика» отражают мое убеждение в том, что для обсуждения этических вопросов одной главы недостаточно. В них рассматривается обоснованное согласие, руководство объединениями участников, корректное проведение опросов и другие темы.

Обычно в книгах, посвященных методам исследований, изложение начинается с простых описательных методов (наблюдение, опрос и т. д.), затем переходит к корреляционным и квазиэкспериментальным и заканчивается методом эксперимента. Без сомнения есть логика в таком расположении материала, но я решил воспользоваться другой схемой. Преподавая этот курс несколько лет назад, я всегда был озадачен тем, что студенты узнают о факторных планах далеко не в начале семестра — кто захочет находить отношения переменных, все еще переваривая ужин, съеденный на День благодарения, — и поэтому решил начать изложение экспериментального метода гораздо раньше, чтобы уделить ему столько времени, сколько потребуется, особенно если у студентов возникнут трудности с пониманием. Кроме того, большинство лабораторных занятий включало экспериментальные планы и я хотел, чтобы студенты имели некоторое представление об исследованиях, проводимых в течение семестра. Такой способ преподавания отражен в организации материала данной книги — эксперимент рассматривается в ней до изложения других методов. Мнения рецензентов разделились — часть из них предпочла бы начать с описательных методов, но большинству понравилась предложенная организация. Я был рад узнать, что некоторые рецензенты, преподающие данный курс и предпочитающие начинать его с описательных методов, тем не менее используют мою книгу, изменив порядок изложения материала. Поэтому стоит отметить, что книга состоит из относительно независимых глав и изменение их последовательности обеспечивает различные варианты преподавания данного курса.

Изменения в третьем издании

Первые два издания книги «Исследование в психологии» были весьма успешными, и мне кажется, что книга получилась неплохой. В третьем издании кроме общей переработки текста были сделаны также следующие дополнения и изменения.

- В ответ на просьбы студентов, я добавил ответы к упражнениям, приведенным в конце каждой главы. В приложение *Е* включена половина из них — остальные упражнения проверяются преподавателем. Ответы на все вопросы можно найти в «Руководстве для преподавателей».
- Проверочные задания типа «вставьте пропущенные слова» в конце главы заменены темами небольших эссе. Это способствует более глубокому пониманию студентами предложенного материала.
- В главе 1 в разделе, посвященном описанию псевдонауки, устаревший пример с биоритмами заменен на более современный пример с бессознательно воспринимаемыми аудиозаписями. В главу, посвященную этике исследований, добавлен новый раздел об особых слоях населения (детях, заключенных).
- Большое влияние на психологию оказал Интернет. Одна из иллюстраций этого — изменение возможностей поиска нужной литературы. Раздел об он-лайновом поиске в базе данных *PsycINFO* (глава 3) переписан в соответствии с новейшими данными.
- Некоторые примеры заменены более интересными и современными (например, изучением так называемого эффекта Моцарта, исследованием бессознательно воспринимаемых аудиозаписей, объединяющих эффекты плацебо и листа ожидания).
- Добавлен материал о качественных процедурах исследования (например, работа с фокус-группами) и о различиях между качественными и количественными исследованиями.
- В раздел, посвященный описательной статистике (глава 4), добавлено описание построения диаграмм методом стебля и листа.
- Обсуждение внешней валидности (глава 5) дополнено рассмотрением кросскультурных факторов.
- Детально разработаны вопросы полевых исследований, создания эквивалентных групп и множественной регрессии.
- Добавлена новая вставка «Этика» с практическими рекомендациями о том, как стать этически компетентным экспериментатором (глава 8), и вставки того же типа, посвященные информированному согласию.
- Разработано описание множественной регрессии (глава 9).
- Обсуждение временных последовательностей расширено изучением различных видов планов (глава 10).
- В главу 11 добавлен пример, иллюстрирующий **план с изменяющимся критерием.**

- Эпилог включает в себя новый раздел, обобщающий знания, полученные студентами в ходе курса («Что я узнал в данном курсе»).
- В приложение *А* добавлен раздел с советами по созданию презентации с использованием рисунков и слайдов.

Педагогические цели

Эта книга имеет несколько особенностей, облегчающих студентам ее изучение. Каждая глава открывается кратким обзором следующей далее информации и постановкой учебных задач. Главы завершаются обобщением важных моментов, заданиями для повторения и упражнениями. В задания входят вопросы с вариантами ответов, из которых нужно выбрать правильный (ответы см. в приложении *Е)*, и темы для небольших эссе. Вопросы направлены не на проверку знания студентами определения изученных понятий, а на способность их применения. Упражнения включают требующие размышлений вопросы и задачи, решения которых находятся с помощью изученных в данной главе понятий. Наиболее важные термины и понятия выделены **жирным шрифтом** и приведены в глоссарии в конце книги.

Благодарности

Данный проект не мог бы начаться, завершиться и увидеть свет в трех изданиях без поддержки многих людей. Особо хочу отметить участие моей жены (Сюзан, корпоративный аудитор, помогала мне концентрироваться на работе, иногда позволяя отвлечься на партию в гольф) и детей (Керри, психолога-экспериментатора и мою коллегу-профессора, и Чарльза, будущего геолога и моего партнера по гольфу). Сотни студентов — слушателей моего курса по методам исследований — вдохновили меня на написание этой книги — до того, как я приступил к работе, многие из них советовали мне перестать жаловаться на существующий учебник и написать свой собственный. Особенно хочу поблагодарить Эми Фасо Райт — лидера студенческой группы, интересующейся когнитивными картами, и основного автора примеров на данную тему (см. приложение *А).* Я очень рад, что она получила докторскую степень по фармакологии.

Я чувствую себя в большом долгу перед Деррилом Брюсом, моим научным руководителем, — когда я учился в Таллахасси, он первым показал мне, насколько увлекательными могут быть психологические исследования, и потом, во время наших (почти) ежегодных трехчасовых завтраков на встречах Американской психологической ассоциации, он был моим наставником. Также хочу поблагодарить моих коллег из «Общества преподавателей психологии» *(Society for the Teachingof Psychology)* (входящего в состав *АРА)* — Вейна Ветина и Стива Дэвиса. В самом начале работы над проектом они смогли убедить меня, что я способен написать книгу, а затем продолжали оказывать мне дружескую помощь и поддержку.

Спасибо рецензентам этого издания за их комментарии и предложения: профессору Вильяму П. Смозермену из Бинхэмптоновского университета *(Binghamton University-SUNT),* профессору Томасу Джойнеру, Университет штата Флорида *(Florida State University),* профессору Джейсону Л. Хику, Университет штата Лу-

изиана, профессору Маргарет Радди, Колледж Нью-Джерси, профессору Дженифер Майерс, Мичиганский университет, профессору Брэдли Х. Смиту, Университет Южной Каролины, профессору Элен Батт, Университет штата Индиана.

В заключение хочу отметить, что редакторы и сотрудники издательства *John Wiley* работали прекрасно, сделав процесс легким, как дуновение ветра (или по крайней мере не таким трудным, как я имел основания ожидать), и угощали обедом на наших совещаниях. Особая благодарность Элен Шац, редактору психологического отдела подразделения *College Division* издательства John *Wiley*. Она оказала большую поддержку в работе, ее замечания всегда были тактичны, а предложения по усовершенствованию книги, казалось, были сделаны человеком гораздо более старшим.

ГЛАВА 1

Научное мышление в психологии

Обзор задач главы

Предмет выбранного вами пособия, возможно, является самым важным из того, что нужно изучить, чтобы стать психологом. В начале этой главы мы хотим показать, что знания о методике исследований — необходимая часть психологического образования. Далее следует вводный обзор различных типов мышления — ненаучного, псевдонаучного и научного. Разграничение науки и псевдонауки для психологии особенно важно, так как многие распространенные «психологические истины» (например, представление об эффективности так называемых «бессознательно воспринимаемых аудиозаписей») суть не что иное, как псевдонаучные заблуждения. Завершается глава разговором о задачах психологии и кратким описанием научной деятельности двух выдающихся ученых, работавших в области экспериментальной психологии, — Элеанор Гибсон и Б. Ф. Скиннера. Изучив данную главу, вы получите знания, которые сделают вас способными:

- доказать важность курса методики исследований;
- объяснить, в чем отличие его главных задач от задач других курсов, предусмотренных психологическим образованием;
- перечислить ненаучные типы мышления и описать их;
- описать характеристики научного типа мышления, на котором основана работа исследователей;
- распознать псевдонаучный тип мышления по его главным особенностям;
- описать основные задачи психологического исследования.

Немецкий психолог Вильгельм Вундт в своем двухтомном труде «Основы физиологической психологии» указал, что его книга представляет собой «попытку обозначить новую *научную* область» (Wundt, 1874/1904; курсив мой. — *Дж. Г.).* Вскоре после выхода книги в печать Вундт основал знаменитую психологическую лабораторию в Лейпциге, привлекшую студентов со всех концов Европы и Америки. Его примеру последовали американские университеты, и к 1892 году было открыто около 20 лабораторий (Sokal, 1992). В том же году была основана Американская психологическая ассоциация. Цель ассоциации была определена следующим образом: «...развитие психологии как *науки.* Все члены ассоциации обязаны принимать участие в этой деятельности» (Cattell, 1895, p. 150; курсив мой. — *Дж. Г.).*

Таким образом, пионеры психологии в Германии и США хотели, чтобы «новая психология» стала наукой. Постепенно психология отделилась от физиологии и философии и стала независимой дисциплиной.

Психологи тех лет видели в новой психологии науку о психической жизни, целью которой является точное понимание того, как психика структурирована, и/или того, как она влияет на адаптацию к окружающей среде. Для научного изучения психики было необходимо разработать особые методы и научить будущих экспериментаторов пользоваться ими, поэтому студенты, изучавшие новую психологию, занимались в лабораториях, где осваивали возможные подходы к исследованию. Одной из самых известных книг по психологии в то время был двухтомный (даже четырехтомный, если учитывать руководство для преподавателей) учебник для лабораторных занятий, написанный в начале XX в. выдающимся психологом-экспериментатором Э. Б. Титченером. Этот учебник использовался до середины 30-х гг.; на нем выросло целое поколение психологов-экспериментаторов (Tweney, 1987).

С тех пор отдельные методы заметно изменились, но и сегодня обучение студентов следует основанной тогда традиции. Таким образом, с самого возникновения психологии как науки преподавание методов исследования было в центре учебного плана. Студент, естественно, может усомниться в том, что он должен изучать курс исследовательских методов просто потому, что «его всегда изучали», поэтому необходимо представить другие, более весомые аргументы.

Зачем нужен этот курс?

Наиболее очевидная причина, по которой стоит заниматься исследовательскими методами, — это возможность таким образом узнать, как эти исследования проводятся. На мой взгляд, события в идеале должны развиваться так: вас заинтересует описание исследования, и вы решите самостоятельно его повторить, приобретая таким образом первый опыт еще будучи студентами, поработав вместе с преподавателями и, возможно, представив результаты на студенческой конференции, вы поступите в аспирантуру, защитите докторскую диссертацию по психологии, начнете карьеру исследователя и в один прекрасный день попадете в число получивших ежегодную награду *APA* за «весомый вклад в науку». Конечно, я реалист и понимаю, что профессиональные интересы большинства психологов лежат не в исследовательской сфере, что многие не станут защищать докторскую диссертацию, а большинство из тех, кто защитит, не станут исследователями, и очень немногие из способных учеников получат престижные награды. Имеет ли смысл заниматься этим курсом, если вы не станете в будущем известными учеными? Конечно, да.

Это следует делать хотя бы по той причине, что курс методики исследований дает прочную основу для освоения других психологических дисциплин (социальной психологии, когнитивной психологии, психологии развития и др.). Различие между курсом методики и другими учебными курсами заключается, по сути, в различии учебного процесса и его содержания. Занятия данным курсом позволят вам освоить саму технологию получения знаний, которую вы будете в дальнейшем

применять к любой области, представленной программами учебного плана. Социально-психологический эксперимент, касающийся конформности, может очень сильно отличаться от когнитивно-психологического исследования кратковременной памяти, но оба эти исследования будут объединены методом — способом получения знаний о предмете. Кроме того, встречающиеся в учебниках примеры исследований станут более понятными, если вы познакомитесь с методами интерпретации поведения, за которым наблюдал ученый.

Чтобы убедиться в этом, перелистайте какой-нибудь учебник по психологии. Очень вероятно, что почти каждый параграф будет содержать рассуждения о каком-либо типе поведения в сопровождении описаний исследований или ссылок на них. У меня на полке, например, стоит книга Майерса (Myers, 1990), в которой описано одно исследование влияния «жесткой» порнографии на мужскую агрессивность (Donnerstein, 1980). Майерс пишет, что экспериментатор «показывал 120 мужчинам нейтральные, эротические и агрессивно-эротические видеозаписи. Затем эти мужчины "учили" других людей (мужчин и женщин) соединять ничего не значащие слоги. Обучаемые выбирали слоги сообразно полученному за неправильный ответ физическому наказанию. Мужчины, которые смотрели фильмы со сценами насилия, применяли к женщинам более сильное наказание» (Myers, 1990, p. 393). Те, кто прочитают это описание, не будучи знакомы с методикой эксперимента, получат лишь общее представление, тогда как знакомые с ней смогут предположить, что исследование проводилось, скорее всего, по факторному плану 2x3, и понять, что оно показало преобладание эффекта взаимодействия над основными эффектами; что обе независимые переменные (характер видеозаписи и пол обучаемого) были управляемыми и это сделало интерпретацию результатов более определенной; а также то, что «жертвы» не испытывали болевых воздействий и были осведомлены о целях исследования (т. е. являлись соучастниками)[1].

При размышлении о дальнейшем ходе эксперимента некоторые задумаются о том, «что произошло бы, если бы между показом видеозаписей и "обучением" прошло больше времени?» или о том, «как бы вели себя в роли "обучающих" испытуемые женского пола?»

Третий довод в пользу занятий экспериментальной психологией заключается в том, что даже если вы не приобретете никаких важных лично для вас знаний в ходе изучения этого курса, знание исследовательских методов сделает вас более разборчивыми и критичными. Мы постоянно сталкиваемся с попытками объяснения поведения. Они поступают из разных источников, от окружающих людей до разнообразных средств массовой информации, как серьезных (статья в солидном журнале о зависимости роста агрессивности от телепередач), так и низкопробных (газета-таблоид). Если последняя большинством не воспринимается всерьез, то написанная профессиональным писателем статья о влиянии телевидения способна вызвать доверие. При этом автор может не иметь никакого представления о важном различии между экспериментальным и корреляционным исследованием, а следовательно, можно ожидать, что корреляционное исследование будет описано в терминах причин и следствий. Изучив главу 9 данной книги, вы легко распознаете такую ошибку. Другой пример — распространенное мнение о том, что загипноти-

[1] К концу данного курса подобный жаргон станет языком вашего повседневного общения.

зированный человек может мысленно вернуться к моменту своего рождения и таким образом получить важную информацию о происхождении своих проблем. Из главы 3 вы узнаете о «скупых» объяснениях, после чего будете с сомнением относиться к подобным утверждениям и задумаетесь о том, какие другие объяснения можно дать рассказам пациентов о своих натальных переживаниях. Вы приобретете скептическое отношение к «бессознательно воспринимаемым» аудиозаписям (подобные записи, скорее всего, продаются в книжном магазине вашего колледжа).

В-четвертых, существует прагматический довод в пользу изучения методов эксперимента. Даже если вы не захотите стать психологами-исследователями, вы, возможно, станете практикующими психологами. Как и исследователю, практикующему психологу необходимо получить научную степень, предпочтительно докторскую. Все будущие консультанты, будущие клинические и школьные психологи будут проводить учебные исследования; данный курс — первый шаг к приобретению необходимых для этого навыков. Кроме того, ваши шансы поступить в хорошую аспирантуру сильно повышаются, если вы а) успешно справились с курсами методики исследований и статистики и б) участвовали в проведении исследований, еще будучи студентами. Исследование Норкросса с соавторами (Norcross, Hanych, and Terranova, 1996) показало, что для поступления в аспирантуру чаще всего требуется знание статистики и методов исследования, тогда как знание других курсов из различных областей психологии (таких, как психология развития или психопатология) не является необходимым во многих программах[1].

Для профессионального психолога исследовательские навыки бесценны. Даже если вы не будете заниматься исследованиями, вам потребуется всегда быть в курсе последних изысканий в вашей области и критически изучать описания проведенных экспериментов. Более того, профессиональная клиническая деятельность подразумевает владение тем же научным способом мышления, который требуется для лабораторных исследований, — при работе с пациентами необходимо выдвигать гипотезы о причинах проблем, проверять их, используя различные терапевтические техники, а также систематически оценивать результаты. Работая в службе социальной помощи, вы будете иметь дело с аттестационными комиссиями и организациями, финансирующими вашу службу, которым потребуются доказательства эффективности вашей профессиональной деятельности. Как вы узнаете из главы 10, профессиональным психологам нередко приходится проводить оценку эффективности своих программ.

Конечно, не все станут профессиональными психологами и получат степень магистра — данный курс поможет будущим бакалаврам выработать навыки, которые потребуются при трудоустройстве. Вы улучшите ваши способности критического и аналитического мышления, научитесь ясно излагать свои мысли в письменной форме, более глубоко овладеете логической аргументацией. Кроме того, вы узнаете, как обобщать и интерпретировать эмпирические данные, искать информацию в библиотеках и электронных базах данных, а также представлять результаты исследований в четкой и понятной форме. Улучшатся также ваши навыки работы с компьютером — вы познакомитесь с программным обеспечением для статисти-

[1] Анализ 1554 аспирантских программ показал, что 85,2% из них «требуют» или «рекомендуют» знание статистики, 66,0% — знание методики, 35,9% — психологии развития/детской психологии и 32,5% — психопатологии.

ческой обработки данных (например, *SPSS)* и разработки способа их представления *(PowerPoint)*.

В завершение отмечу, что по прохождении курса исследовательских методов вы усвоите определенный образ мышления. Как было отмечено ранее, другие курсы направлены на овладение содержанием знаний в определенных областях психологии и направляют внимание на то, что известно о явлении *X*. Данный курс сфокусирован на процессе, с помощью которого приобретается знание об этом *X*. В основе этого процесса лежит научный тип мышления. Но прежде, чем рассказать об особенностях научного мышления, необходимо описать распространенные способы познания окружающего мира.

Способы познания

Давайте задумаемся над чем-нибудь, что мы считаем безусловно верным. Это может быть нечто совсем простое, вроде мысли о том, что сад лучше поливать утром, а не вечером, или нечто сложное, как, например, религиозное чувство. Как мы приходим к уверенности в чем-то?

В 1877 г. американский философ-прагматик Чарльз Пирс в своей работе «Становление убеждений» *(The Fixation of Belief)* описал четыре пути формирования убеждений — упорство *(tenacity),* авторитет, метод априори и научные методы. Пирс считал, что первые три подхода широко распространены, но несостоятельны и что только наука является истинным способом получения знаний.

Упорство

Первый путь, описанный Пирсом, **упорство** — это рожденная страхом перед неопределенностью безусловная вера в правильность некоторого суждения, не поддающаяся никакой критике. Люди, таким образом строящие свои убеждения, обращают внимание только на подтверждающую информацию и нечувствительны к опровергающим свидетельствам. Пирс не описывает, как подобные убеждения возникают, отмечая лишь, что они весьма примитивны и рождаются из «инстинктивных неприятия или неуверенности» (Tomas, 1957, p. 15). Вероятно, такие убеждения формируются, когда человек многократно слышит некую «истину» и не встречается с противоречащими ей фактами. Например, студенты колледжей 60-х гг. верили в существование конфликта между поколениями и как мантру повторяли слова: «Не доверяй никому старше 30» (сейчас этим людям около 50, и они с большим недоверием относятся ко всем, кто моложе 30).

Убеждения, которых упорно придерживаются, дают своим носителям ощущение комфорта, внося определенность в мир, полный сомнений. Однако на их формирование серьезно влияют личные предубеждения. Современные социальные психологи отождествляют понятие упорства с понятием **ригидности мнения — упрямым** нежеланием рассматривать факты, противоречащие избранной точке зрения (Anderson, Lepper & Ross, 1980). Упрямый человек, «знающий», что в 1947 г. американский самолет сбил летающую тарелку и доставил погибших инопланетян в город Росвел в штате Нью-Мексико, проигнорирует любые научные опровержения этого факта (или воспримет их как намеренную ложь). Феномен упорства

проявляется в силе предубеждений — люди с сильными предубеждениями упорно верят в свою правоту, и переубедить их практически невозможно. На убежденного в том, что получающие пособие по безработице просто не хотят работать, издеваются над системой и имеют дорогие машины, никак не подействует опыт социологических исследований о причинах возникновения этого стереотипа.

Авторитет

Второй путь «закрепления убеждений», описанный Пирсом, — это **авторитет.** При 1
объяснении этого явления Пирс использовал пример властей штата, постоянно повторяющих одни и те же «правильные» вещи и стремящихся иметь в глазах населения вид принципиальных (Tomas, 1957, p. 16). Влияние авторитета осуществляется по-разному: будучи детьми, мы верим тому, что говорят нам наши родители (хотя бы некоторое время); став студентами, мы признаем авторитет преподавателей и верим написанному в учебниках; как пациенты, мы принимаем прописанные лекарства и верим, что они помогут, и т. д. Несомненно, полагаясь на мнение авторитетов при формировании собственных убеждений, мы не думаем о том, что авторитеты могут заблуждаться. Родители часто передают детям предвзятые мнения, преподаватели иногда ошибаются, а их знания могут быть неполными или отражать личные предпочтения, учебники могут содержать ошибки и недостоверную информацию, врач может неверно поставить диагноз или прописать неподходящее лекарство, а тот факт, что правительство часто принимает неверные решения, вряд ли требует комментариев.

С другой стороны, из авторитетных источников, а в особенности от экспертов в своей области, мы узнаем многие важные вещи. Например, мы читаем «Консьюмер рипортс», смотрим прогноз погоды и прислушиваемся (иногда) к словам врачей о необходимости физической активности и правильного питания. Кроме того, безусловно авторитетны гении литературы и искусства, объясняющие нам многое в человеческом поведении. Как можно читать Шекспира, Диккенса или Остин не понимая при этом, что проникаешь в сущность человеческой природы?

Рассуждение и логика

Пирс считал, что упорство и опора на авторитет имеют серьезные недостатки как пути познания. Они приносят успокоение и иногда случайно приводят к постижению истины, но при этом очень часто ведут к ошибкам. Третий путь «закрепления убеждений» является шагом вперед по сравнению с предыдущими, хотя и имеет свои недостатки. Пирс описал его как «постепенное формирование убеждений в гармонии с естественными законами причины и следствия» (Tomas, 1957, p. 20) в результате совместного обсуждения проблемы с людьми, придерживающимися различных взглядов. Убеждения, образованные таким способом, называются «принятыми на основании здравого смысла» (p. 20). Таким образом, третий подход заключается в авторитете разума для достижения согласия между людьми, которые взвешивают достоинства того или иного мнения. Пирс назвал этот подход **методом априори,** так как он основывается на логике и аргументировании, а не на непосредственном опыте. Мнение логически выводится из сделанного ранее допущения (априори в переводе с латинского означает «заранее, без проверки»). Пирс

с легкой тенью сарказма писал, что философы-метафизики любили метод априори и одни из них красноречиво доказывали истинность своих убеждений, а другие не менее красноречиво опровергали их и доказывали нечто противоположное (например, дискутируя на тему, является ли тело и разум единой или же двумя раздельными сущностями). В результате убеждения оказываются зависимыми от моды. Говоря о метафизике, в частности, Пирс писал, что «маятник с давних пор качается туда и обратно от более материалистической к более идеалистической философии» (Tomas, 1957, p. 24).

Опыт

Еще один важный вид познания, описанный Пирсом, — это **эмпиризм** — получение знаний с помощью непосредственного наблюдения или опыта. Вскоре вы узнаете, что постановка «эмпирических вопросов» является важной частью научного мышления и что поговорка «опыт — лучший учитель» абсолютно верна. Конечно, опасно в поисках истины слепо полагаться на чей-то опыт, но проблема в том, что наш личный опыт вынужденно ограничен. Кроме того, на его интерпретацию воздействует множество факторов — феномен, названный социальными психологами «искажением, вызванным социальным ожиданием». Ригидность мнения, упомянутая в связи с понятием упорства, — одно из таких искажений. Другие его разновидности — переоценка по причине запоминаемости и искажение через подтверждение. **Переоценка по причине запоминаемости** происходит, когда мы переживаем необычные или очень яркие события, а затем переоцениваем их возможную частоту (Tversky & Kahneman, 1973). Так, люди, которые часто смотрят по телевидению передачи о преступниках, переоценивают вероятность стать жертвой преступления. Или, например, оттого, что крушения самолетов привлекают большее внимание средств массовой информации, чем автокатастрофы, и потрясают аудиторию, некоторые люди не верят, что авиатранспорт гораздо менее опасен, чем автомобильные поездки. **Искажение** по **подтверждению** — это тенденция к поиску информации, подтверждающей имеющееся мнение, и к пренебрежению опровергающими его фактами (Wason & Johnson-Laird, 1972). Например, люди, которые верят в экстрасенсорное восприятие, будут внимательны к тем моментам, когда они «думали о мамочке, как вдруг зазвонил телефон, и это была она!» и проигнорируют многочисленные случаи, когда они а) думали о мамочке и она не позвонила и б) не думали о мамочке и она позвонила. Им также не придет в голову, что если они беседуют с матерью примерно раз в две недели, количество мыслей о ней должно будет увеличиться к концу двухнедельного периода, а следовательно, будет расти вероятность «попадания». Опыт может быть надежным штурманом в море жизненных трудностей, но необходимо помнить и о его ограниченности.

Переоценка по причине запоминаемости и искажение по подтверждению могут заметно влиять на убеждения, сформированные опытом. Ярким примером является то, что происходит при исправлении студентами ответов, данных в тесте на выбор из нескольких вариантов. Многие верят, что при этом обычно правильный ответ исправляется на неправильный. Такое событие очень хорошо запоминается, возможно, потому, что сказывается на полученной оценке. Когда убеждение начало формироваться, оно подкрепляется каждый раз, когда делается подобное ис-

правление. Кроме того, студенты склонны не обращать внимания на случаи, когда они меняют один неправильный ответ на другой или неправильный ответ на правильный. Запоминается только ситуация замены правильного ответа на неправильный, снижающая оценку за работу. Это убеждение опровергается исследованием, показавшим, что наиболее часто (около 58% случаев) происходит замена неправильного ответа на правильный. Обратное бывает только в 20% случаев, а оставшиеся 22% приходятся на случаи замены одного неверного ответа на другой (Benjamin, Cavell & Shallenberger, 1984).

Способы познания и наука

По Пирсу, самым надежным способом формирования убеждений являются научные методы, обеспечивающие познание «явлений как они есть, вне зависимости от нашего мнения о них» (Tomas, 1957, стр. 25). Таким образом, по мнению Пирса, основным достоинством науки является объективность, которую он противопоставляет субъективности, причем объективность означает полное отсутствие предубеждений и психологических искажений. Прежде чем подробно обсуждать научное мышление, необходимо отметить, что ученые — это такие же люди, как и все остальные, они бывают склонны к упорству, полагаются на авторитеты, спорят друг с другом в априорном стиле и учатся на личном опыте.

Что касается упорства, то иногда ученые придерживаются излюбленной теории или методологии и спустя долгое время после того, как все остальные забудут ее. Также они бывают весьма далеки от желания воспринять новую идею. Чарльз Дарвин как-то полушутя написал, что смерть в 60 лет могла бы стать неплохим выходом для ученых, так как после вступления в этот возраст они «наверняка будут возражать против всех новых учений» (Boorstin, 1985, p. 486). С другой стороны, историк науки Томас Кун в 1970 г. отметил, что отказ от принятия теории, подтвержденной лишь несколькими экспериментами, сомнения в ее валидности могут способствовать тщательной проверке этой теории. Решительный (упорный) защитник теории может помочь выработать все ее ресурсы, прежде чем она будет забыта в научном обществе. Процесс оценки, развития и заката теорий рассматривается в главе 3.

Ученые (и психологи-исследователи в их числе) могут также оказаться восприимчивыми к влиянию авторитета. «Авторитетами» обычно являются другие ученые, вероятность достоверности суждений которых довольно высока. Однако исследователи не склонны принимать нечто за истину только потому, что известный ученый считает это верным. Обычно ученые поступают в соответствии с девизом, помещенным на входе в здание Британского королевского общества, — *«Nullius in Verba»,* согласно которому следует «не доверять чужим словам и во всем убеждаться на практике» (Boorstin, 1985, p. 394). К сожалению, как было отмечено ранее, процесс «убеждения на практике» рождает искажения, вызванные социальным ожиданием.

Метод априори часто применяется в науке, когда ученые спорят друг с другом и пытаются прийти к единодушному мнению по некоторому вопросу, нередко безрезультатно (например, обсуждая, можно ли считать подходящей метафорой для мозга слово «компьютер»). Как вы узнаете из главы 3, при разработке идей исследо-

вания и оценке его выводов ученые опираются на законы логики и индуктивные/дедуктивные умозаключения. Но хотя научное мышление включает в себя элементы описанных выше способов познания, оно имеет также особые, присущие только ему черты. Описанием этих черт мы и займемся в следующем разделе.

Особенности научного мышления в психологии

Способ мышления, характерный для ученых в целом и исследователей-психологов в частности, имеет определенные особенности. Во-первых, исследователи полагают, что явления действительности, в том числе человеческое поведение, протекают не беспорядочно, а по определенным законам, следовательно, они упорядочены, и могут быть предсказаны с некоторой степенью вероятности. Таким образом, делаются два предположения — **детерминизма** и **возможности открытия:** все события имеют причины, и эти причины могут быть открыты с помощью научных методов. Это не означает, однако, что события могут быть предсказаны с вероятностью 100%, это говорит лишь о том, что психологические явления возникают не случайно, а в соответствии с некоторыми закономерностями и эти закономерности могут быть открыты. Давайте рассмотрим подробнее теорию детерминизма.

Детерминизм

Студенты часто бывают сбиты с толку, прочитав, что психологи считают человеческое поведение «детерминированным». Они полагают, что это означает «предопределенный» или «предустановленный», но это неверно. Человек, верящий в абсолютное предопределение, думает, что каждое событие заранее предрешено, например Богом. Он считает, что человек должен принимать жизнь такой, как она есть, и не может ничего изменить в ней. В противоположность этому традиционная теория детерминизма утверждает лишь, что все события имеют причины. Некоторые философы защищают строгий детерминизм, согласно которому наличие причинно-следственной организации Вселенной дает возможность предсказания событий со 100% вероятностью, хотя бы теоретически. У других философов под влиянием открытий физики XX в. сформировались более умеренные взгляды — они являются сторонниками концепции так называемого вероятностного, или **статистического, детерминизма** и считают, что события могут быть предсказаны с вероятностью выше вероятности случая, но ниже 100%. Большинство психологов-исследователей придерживается этой теории.

Несмотря на признание ограничения возможности предсказания порогом вероятности менее 100%, теория детерминизма все еще вызывает вопросы — складывается впечатление, что она противоречит идее свободной воли. Если каждое событие вызывается некоторой причиной, то как можно свободно выбрать одно направление действий и пренебречь другим? Психологи на это могли бы ответить вопросом: как вообще возможно узнать что-либо о поведении, если теория детерминизма неверна? Представьте себе мир, в котором поведение человека абсолютно непредсказуемо. Как вы смогли бы выбрать, выходить замуж за Эда или за Теда? Как бы вы решили, посещать или нет курс профессора Джонс?

Конечно, на поведение влияют многочисленные факторы, и поэтому очень трудно с уверенностью сказать, что будет делать конкретный человек в определенный

момент времени, но тем не менее поведение строится по определенным моделям, а следовательно, предсказуемо. Например, зная, что дети стараются делать то, что однажды уже «сработало», несложно предсказать капризное поведение ребенка в отделе игрушек наполненного людьми магазина, если раньше это помогло ему получить игрушку. А так как новый образец поведения, использованный в одной ситуации, имеет тенденцию к «генерализации» в похожих ситуациях, можно предсказать каприз ребенка в следующем магазине, если ранее каприз «сработал» в другом.

По поводу свободного выбора философ науки и позитивист Рудольф Карнап отмечает, что без детерминизма он не имеет смысла, так как выбор должен делаться на разумных основаниях, а такие основания есть только в мире, содержащем законы. Согласно Карнапу, без «постоянства отношений причин и следствий... вообще невозможно сделать свободный выбор. Выбор предполагает обдуманное предпочтение одного образа действий перед другим. Как можно сделать выбор, не предвидя последствия этих действий?» (Carnap, 1966, p. 220). Итак, Карнап считает, что идея свободного выбора без идеи детерминизма не имеет смысла. Таким образом, выбор между Эдом и Тедом осмыслен, только если вы можете предсказать некоторые особенности их поведения (например, Эд будет более надежным партнером). Решение о посещении занятий профессора Джонс может основываться на ее репутации преподавателя, от которого можно ожидать справедливого отношения к студентам.

Большинство психологов-исследователей считает, что вопрос о существовании свободной воли не может быть решен наукой. Свободно или нет мы делаем выбор в нашей жизни — это вопрос философский, и к выводу о свободе воли мы должны прийти путем самостоятельных размышлений (и возможно, не без помощи упорства и/или авторитета). Однако исследователям ясно, что выбор имеет смысл для людей, только если события в мире предсказуемы. Поэтому когда психолог изучает поведение и находит в нем закономерности, этот факт не исключает и даже не ограничивает возможность человеческой свободы. Действительно, если Карнап прав, то исследование психолога может только усилить нашу способность выбора, так как оно увеличивает знание о границах наших возможностей.

Объективность

Вторая, и наиболее привлекательная, по мнению Пирса, особенность научного мышления — это его относительная **объективность.** Согласно Пирсу, быть объективным означает избегать таких человеческих факторов, как ожидания и склонности. В соответствии с такими взглядами объективный ученый представляется подобным машине в беспристрастном поиске истины. Сегодня, однако, никто не думает, что ученый может полностью отвлечься от существующих мнений, а быть объективным не значит лишиться естественных человеческих особенностей. Объективное наблюдение — это наблюдение, подтверждаемое несколькими исследователями. Для научного мира это означает необходимость достаточно точных определения терминов и описания процедуры исследования, для того чтобы другие ученые могли повторить данное исследование и получить аналогичные результаты. Такой процесс повторения исследования для проверки его результатов называется «репликация» (см. главу 7). Если результаты совпадают с полученными ранее, то ве-

роятность существования данного психологического феномена возрастает, в противном случае — снижается. Последовательно неподтверждаемые результаты в конечном счете не принимаются — это делает науку самокорректирующейся системой. Как вы узнаете из следующей главы, отсутствие подтверждения результатов позволяет также раскрыть научное мошенничество.

Чтобы повторить исследование, необходимо точно знать, что именно было сделано при оригинальном изыскании. Для описания исследовательского проекта существует особый набор правил, подробно изложенный в «Общем руководстве Американской психологической ассоциации» (Manual of the American Psychological Association, 1994) — неоценимом источнике информации для всех, кто составляет отчет о результатах исследования или любой другой документ. Приложение *А* представляет собой пособие по написанию лабораторного отчета в соответствии со стандартами *APA.* Оно создано на основе данного руководства и является хорошим введением для обучающегося составлению отчетов.

Объективность была проблемой психологической науки с самого ее появления. Новая наука-психология была определена как «наука о психической жизни», а один из ее самых первых методов был назван **интроспекцией.** Процедура в значительной степени варьировалась от лаборатории к лаборатории, но в целом интроспекция была особой формой самоотчета. Участники эксперимента выполняли задание, а затем описывали полученный при этом сознательный опыт. Некоторое представление об интроспекции дает вставка 1.1, содержащая дословное интроспективное описание эксперимента по изучению внимания. Она показывает, что интроспективный способ мышления был частью бытового познания психологов того времени.

Проблема интроспекции как метода заключалась в том, что хотя участники проходили серьезное обучение, чтобы устранить возможность искажений при самонаблюдении, этот метод был субъективным по своей природе (вы не можете подтвердить истинность моей интроспекции, а я не могу подтвердить истинность вашей). Данная проблема заставила Джона Б. Уотсона и других психологов принять следующее решение: для того, чтобы психология действительно стала «наукой», необходимо научиться измерять то, что наблюдается непосредственно (двумя или более наблюдателями), и таким образом может быть объективно подтверждено. Уотсон считал, что поведение как раз подходит для этого. Его решительное утверждение о том, что психология должна основываться на возможности наблюдения и измерения данных, принесло ему титул «основателя бихевиоризма» как научного направления. Сегодня термин «поведение» входит в определение психологии в любом учебнике по введению в психологию.

Современные психологи принимают в качестве данных поведение и при исследовании внимания не используют подробные интроспективные отчеты в отличие от Далленбаха (см. вставку 1.1). Они разрабатывают эксперименты, в ходе которых вывод о характере внимания делается на основании наблюдаемого поведения, например на основании количества допущенных следящим за ударами метронома испытуемым ошибок при сложении чисел. Также, скорее всего, такой эксперимент будут наблюдать два независимых исследователя, которые затем придут к общему мнению.

ВСТАВКА 1.1

История — знакомство с интроспекцией

Нижеследующий интроспективный отчет был сделан при проведенном Карлом Далленбахом в 1913 г. исследовании внимания. Участников попросили одновременно прослушать удары двух метрономов, работающих с разной скоростью, и сосчитать количество ударов, приходящихся на промежуток времени между двумя совпадающими ударами (т. е. между случаями, когда удары обоих метрономов происходят одновременно). Кроме того, их попросили в процессе подсчета выполнять другие задания, например непрерывно складывать числа вслух. Излишне говорить, что эти задания были направлены на выявление пределов внимания. По окончании задания один участник написал: «Звуки метрономов в виде серий отрывистых щелчков были ясно осознаны только четыре или пять раз за время эксперимента и сначала очень мешали. Они сопровождались ощущением напряжения и неприятным чувством. Далее мое внимание было направлено на сложение чисел. Они воспринимались на слух или зрительно и иногда представляли собой изображения на темно-сером фоне, расположенные в трех футах прямо передо мной... Когда этот процесс ясно осознавался, звуки метрономов были очень смутными и едва различимыми» (Dallenbach, 1913, p. 467).

Обратите внимание, что участник эксперимента старался описать все явления, возникающие в сознании во время выполнения заданий, в том числе сенсорные (напряжение), эмоциональные (неприятное чувство) и образные, звуковые и зрительные. Затруднения с одинаково «ясным осознанием» нескольких заданий привели Далленбаха к выводу об ограниченности внимания. Данное явление было уточнено современными учеными и определено как «избирательность» внимания (Broadbent, 1985).

Научное мышление не прекращается, когда ученый покидает лабораторию. Во времена расцвета интроспективного метода психологи часто думали в терминах этого подхода, и в письмах рассказывали друг другу о последних событиях, описывая свои сознательные переживания. Например, в письме в Корнелл Е. Б. Титченеру Эдмунд Сенфорд из Университета Кларка описывает сильную грозу, используя интроспекцию. Сенфорд пишет, что он много наблюдал этим летом и не нашел в этом ничего кроме ощущений от органических процессов и других, неприятно окрашенных. С когнитивной стороны - спазм апперцепции небольшой группки идей, связанных с вещами пугающими и вызывающими инстинктивные действия и мысли... Когда гроза должна была вот-вот начаться, наблюдались различные кардиальные и висцеральные симптомы. Когда она началась, симптомы приняли обычный характер и перестали быть столь заметны как при ожидании грозы (Sanford, 1910).

Информационный подход

Третья особенность научного мышления в психологии — **информационный подход** к исследованиям. Это означает, что выводы о поведении должны основываться на объективной информации, полученной научными методами. Например, утверждение директора приемной комиссии колледжа о том, что «поступившие в этом году подготовлены лучше, чем те, кто поступал в прошлые годы» (в некоторых школах это говорят каждый год), заставит мыслящего научно человека ответить: «Давайте посмотрим данные по этому вопросу за этот и предыдущие годы»,

а также спросить: «Что вы подразумеваете под лучшей подготовкой?» Кроме того исследователи стараются выяснить, насколько информация, приведенная в поддержку определенного утверждения, соответствует этому утверждению. Если кто-то утверждает, что постоянный просмотр телевизора снижает детскую креативность, то ученые сразу спросят о типе и количестве полученных данных (например, как измерялась креативность и сколько собрано информации), о процедуре, используемой для сбора данных, и о способе их статистического анализа.

Исследователи-психологи используют информационный подход даже в повседневной жизни — они нередко размышляют о том, что говорят объективные данные по поводу какой-либо темы. Даже замечание соседа о том, что помидоры в этом году зреют лучше, чем раньше, может вызвать у исследователя множество вопросов для проверки этого утверждения (как именно вы подсчитывали помидоры последние два года? вы подсчитывали количество помидоров, собранных в день, или количество созревших в день? как вы определяете «зрелость»?). Есть определенный риск в использовании этого подхода — соседи, возможно, станут вас избегать.

Крайнюю степень интереса к сбору информации можно увидеть в примерах из жизни сэра Фрэнсиса Гальтона, британского исследователя, жившего в XIX в. Круг его научных увлечений включал дисциплины от географии и метеорологии до психологии. О важности его открытий в области тестирования умственных способностей будет рассказано в главе 9. Гальтон был увлечен построением выводов на основании собранных данных. Так, однажды он измерил интерес публики к различным театральным постановкам, подсчитав количество зевающих во время спектакля. Он изучал механизм ассоциации, подсчитывая количество родственных идей, приходящих ему в голову во время утренней прогулки, а также собирал данные о возрастной потере слуха — для выявления характеристик слуха он изобрел особый прибор (свисток Гальтона), издающий звуки разной высоты (Galton, 1883). Наиболее необычной из попыток Гальтона сделать выводы на основании информационного подхода было его изучение «эффективности молитвы» (Galton, 1872). Как и его родственник Чарльз Дарвин, Гальтон скептически относился к религии и решил проверить, «работают» ли молитвы. Гальтон рассудил, что если молитвы действительно эффективны, то молящиеся больные должны выздоравливать быстрее остальных. Аналогично люди, которые много молятся (например, духовенство) или за которых много молятся (например, король и королева Англии), должны жить дольше обычных людей. Гальтон также предположил, что шансов затонуть у кораблей с миссионерами на борту (за которых молятся многочисленные оставленные дома родственники) гораздо меньше, чем у других кораблей. Ни одно их этих предположений не сбылось. Например, работая с биографическими словарями, он нашел, что большее количество служителей религии живет в среднем 66,42 лет, тогда как адвокаты (люди, имеющие мало шансов на молитвенную поддержку) живут примерно столько же — 66,51 лет (данные из Forrest, 1974, p. 112). Гальтона по понятным причинам критиковали за его упрощенное понимание назначения молитвы, а его статье о молитве было три раза отказано в печати на том основании, что она «слишком убедительна и оскорбительна, чтобы не растревожить осиное гнездо» (цитируется по Forrest, 1974, p. 111), но его исследования несомненно иллюстрируют информационный подход к мышлению.

Эмпирические вопросы

Как было отмечено ранее, эмпирический способ — это способ познания с помощью непосредственного наблюдения или личного опыта. **Эмпирические вопросы** предполагают ответы на основании систематического наблюдения или опыта как характерных особенностей научной методологии. Такие вопросы должны быть достаточно строго сформулированы, чтобы при ответе на них можно было сделать некоторые предсказания. Из главы 3 вы узнаете, что формулировка вопросов представляет собой начальный этап любого исследования, а также узнаете, как сформулировать хороший эмпирический вопрос.

Формирование понятия об эмпирическом вопросе мы начнем со сравнения его с вопросами, на которые нельзя ответить, основываясь на эмпирике. Вспомните, например, вопрос об отношении души и тела, приведенный Пирсом для иллюстрации описанного им метода априори. Философы долгие годы спорили по этому поводу (и спорят сегодня), и, по мнению Пирса, маловероятно, что этот вопрос когда-либо будет решен. Вопрос о том, являются ли тело и душа (психика) двумя разными сущностями или одной, просто не является эмпирическим. На данную тему можно задать множество эмпирических вопросов. Чтобы узнать о влиянии умственной активности (психика) на физическое здоровье (тело), можно спросить: «Как действует психологический стресс на иммунную систему?», чтобы уточнить влияние тела на психическое состояние: «Как физическая усталость сказывается на способности к решению задач?»

Хотя психологи-исследователи считают научный подход идеальным для получения ответов на вопросы, в нашей жизни есть вопросы, ответить на которые ученые не могут, — например, существует ли Бог, или что лежит в основе природы человека, добро или зло. Это важные вопросы, но научно ответить на них невозможно. Конечно, можно эмпирически исследовать факторы, способствующие вере в Бога или совершению хороших или дурных поступков. Так можно поставить следующие эмпирические вопросы:

- Усиливается ли вера в Бога с возрастом (т. е. с приближением смерти)?
- Меньше ли помогают окружающим, если затраты на помощь превышают личную пользу от нее?

Обобщая сказанное, я бы назвал психологов-исследователей «скептическими оптимистами». Они открыты новым идеям и оптимистично относятся к применению научных методов для проверки этих идей, но в то же время они весьма несговорчивы и не признают утверждений, не основанных на опытных данных. Исследователи постоянно размышляют о новых способах научной проверки идей, они уверены в том, что истина откроется тем, кто ставит эмпирические вопросы и отвечает на них, и готовы (иногда неохотно) изменить свои убеждения, если ответы на эти вопросы не соответствуют их ожиданиям.

Психологическая наука и псевдонаука

Все люди интересуются особенностями человеческого поведения и высказывают свое мнение по поводу его причин и внутренних механизмов. Многие из таких утверждений основываются на серьезных научных исследованиях, проводимых по прави-

лам, которые вы узнаете, прочитав данную книгу, и на только что описанном «скептическом оптимизме». Это означает, что большое количество наших знаний базируется на способах мышления и методах, характерных для истинной науки. Но кроме того, многие утверждения делаются от имени психологической науки, имея в основе способы мышления и методы, не истинно научные, а псевдонаучные (префикс «псевдо» в переводе с греческого означает «ложный»). Термин **псевдонаука** используется для характеристики исследований, которые на первый взгляд проводятся научными методами, но на самом деле базируются на неадекватной, ненаучной основе и, как правило, приводят к ложным выводам. Карикатура Сидни Харрисона (рис. 1.1) иллюстрирует печальную правду об отношении людей к псевдонауке — ее популярность. Привлекательность псевдонауки подтверждается распространенностью телефонных служб предсказателей и большим количеством веб-сайтов, посвященных сомнительным психическим феноменам. Как же отличить истинную науку от псевдонауки и чем объясняется высокая популярность последней?



Рис. 1.1. Незаслуженная популярность псевдонауки

Как распознать псевдонауку?

В конце XIX в. в нью-йоркской фирме Фаулера и Уэллса можно было заказать изображения «Символической головы и френологической карты» за 10 центов, а за 1 доллар 25 центов к этому прилагалась инструкция «Как читать характер: новое иллюстрированное руководство по френологии и физиогномике» (Anonymous Advertisment, 1881). Эти материалы предлагали «научное» определение характера по форме черепа.

В конце XX в. за 30 долларов на многих веб-сайтах принимались заказы на «бессознательно воспринимаемые» аудиозаписи, способные улучшить вашу жизнь. Просто прослушав такую запись, можно было значительно развить память, сбросить вес, бросить курить, увеличить потенцию и чувство собственного достоинства, разбогатеть, играя на фондовом рынке, или даже стать более искусным охотником на оленей (buckblaster.com). Эти записи позволяли решать любые жизненные проблемы с минимальными усилиями («Просто расслабьтесь и включите кассету — остальное за вас сделаем мы»).

Как показывают эти примеры, люди готовы платить за самопознание и саморазвитие, особенно если это делается с помощью научных методов, которые легко понять и применить. И френология, и индустрия «бессознательно воспринимаемых аудиозаписей» относятся к псевдонаучной области. Конечно, между ними есть существенные различия, но в целом оба примера демонстрируют основные черты псевдонауки.

Наукообразие

Люди, занимающиеся псевдонаукой, всячески стараются связать ее с настоящей наукой или создать видимость такой связи. Иногда первичным источником псевдонауки действительно является истинная наука, в других случаях ее положения имеют такой вид, что их легко спутать с научными. Френология иллюстрирует первый случай, «бессознательно воспринимаемые записи» — второй. Френология берет свое начало в попытках ученых показать, что различные участки мозга выполняют различные функции, и ее можно рассматривать как одну из первых теорий локализации функций мозга (Bakan, 1966). Она была создана в конце XVIII в. венским врачом Францем-Иосифом Галлем, подтвердившим принцип контрлатеральности мозга (левая часть мозга управляет правой половиной тела, и наоборот). Название «френология» получила его теория о наличии взаимосвязи между особенностями личности и структурой мозга. Она предполагает, что а) различные личностные и интеллектуальные свойства (способности) связаны с различными участками мозга (рис. 1.2), б) сильным способностям соответствуют относительно большие по размеру участки и в) по форме черепа можно сделать выводы об относительной силе способностей. Соответственно предполагалось, что по форме черепа можно определить свойства личности и уровень интеллекта. Несмотря на то что в середине 1800-х гг. положения френологии были опровергнуты блестящими исследованиями французского врача Пьера Флорана, она оставалась популярной до начала XX в. (вставка 1.2). Во второй половине XIX в. френологический бизнес процветал, хотя сама френология была отвергнута учеными — функционировали френологические общества, выпускались популярные журналы, а френологический анализ применялся в любых жизненных ситуациях, от выбора профессии до подбора честного слуги. Даже отвергнутая наукой теория может найти отклик в душе людей. Для психологии это большая проблема, так как практически любая теория, касающаяся человеческого поведения, обладает известной привлекательностью и находит своих сторонников. Новейшим подтверждением этого является продажа «бессознательно воспринимаемых аудиозаписей».

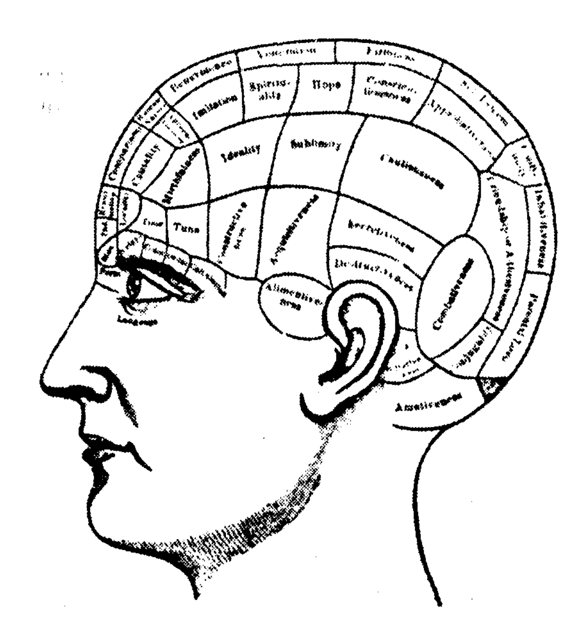


Рис. 1.2. Голова, отображающая схему локализации различных способностей, предложенную френологами

ВСТАВКА 1.2

Классические исследования - опровержение френологии

В 1846 г. вышла в свет книга, озаглавленная «Исследование френологии» *(Phrenology Examined).* Ее автором был Пьер Флоран (1794-1867) - известный французский физиолог и хирург, установивший роль полукружных каналов уха в поддержании состояния равновесия, открывший локализацию дыхательного центра в продолговатом мозге и обнаруживший анестезирующие свойства хлороформа (Kruta, 1972). Он был злейшим врагом френологии и не пытался смягчить смысл слов, когда писал: «Все учение [френологии] состоит из двух фундаментальных утверждений: о том, что процесс постижения действительности зависит от мозга и о том, что каждая конкретная способность постижения имеет в мозге представительство в виде определенного органа. Из этих двух утверждений первое определенно не содержит ничего нового, а второе скорее всего ничего истинного» (Flourens, 1846/1978, p. 18)

Чтобы опровергнуть утверждения френологов, Флоран провел экспериментальное исследование проблемы методом «абляции». Хотя процедура была разработана еще до Флорана, введенные им усовершенствования прочно связали имя исследователя с названием метода. Не прибегая к естественному эксперименту, возникающему в ситуации случайного поражения мозга, Флоран удалял отдельные участки мозга и наблюдал последствия удаления (абляции, от лат. *ablatio* - удаление). Если в результате удаления теряется возможность видеть, то вероятно, что соответствующий участок мозга каким-то образом связан со зрением. Понятно, что метод требовал использования в качестве объекта исследова-

ния животных - Флоран проводил эксперименты с различными видами, например, с собаками и голубями.

Задача Флорана заключалась в демонстрации того, что определенные участки мозга, предположительно выполняющие функцию *X,* в действительности выполняют функцию У, а также того, что кора мозга представляет собой единое целое, а не сумму отдельных функций, локализованных в разных участках. Одним из направлений исследования было изучение мозжечка. По мнению френологов, этот участок мозга отвечал за сексуальную активность и являлся центром «влюбчивости». Например, Иоганн Г. Шпурцхайм (1832-1878) в своей работе «Очерки по френологии» утверждал, что сексуальность «возникает по мере развития этого участка и зависит от его размера. У детей, например, мозжечок меньше, чем у взрослых, а у женщин и самок животных он меньше, чем у мужчин и самцов животных» (р. 28). Очевидно имея в виду некоторые единичные данные, Шпурцхайм указал, что иногда мозжечок «достигает больших размеров у детей, и тогда его особая функция, порождающая определенные склонности, проявляется в раннем возрасте» (р. 28).

Флоран не обнаружил подобной зависимости. Он высмеял порочный круг, возникающий при обозначении особенности поведения как «способности» и последующем объяснении этого поведения в терминах способностей: «Что это за философия, пытающаяся объяснить явления действительности с помощью слова? Вы наблюдаете... склонность у животного ... вкус или талант у человека, затем каждую особенность просто обозначаете как способность и думаете, что тем самым явление объяснено. Но задумайтесь, ваша *способность* - это не более чем *слово,* это просто название факта, а все трудности [его объяснения] никуда не исчезли» (Flourens, 1846/1978, р. 39; курсив автора).

Показать отсутствие связи между сексуальной мотивацией и мозжечком не составило труда. Осторожно удаляя отдельные участки мозжечка, Флоран продемонстрировал, что он является центром координации движения. Например, голуби с удаленным мозжечком не могли координировать движение крыльев для полета, а собаки не могли правильно ходить - они шатались, падали и натыкались на окружающие предметы, тогда как раньше у них не возникало подобных проблем. Изменение сексуальной мотивации выявлено не было, хотя стоит отметить, что поражение мозжечка наверняка вызовет затруднение копуляции. Также Флоран установил, что степень нарушения движения прямо пропорциональна размеру удаленной области мозжечка.

Флоран также проводил исследования по удалению коры мозга и обнаружил аналогичную взаимосвязь между размером удаленной зоны и степенью возникающих нарушений. Он не нашел признаков локализации функций в отдельных участках коры и заключил, что кора действует как единое целое и отвечает за восприятие, мышление и волю в целом. Например, голуби, у которых удалена кора мозга или большая ее часть, были способны воспринимать окружающую действительность, но не демонстрировали признаков понимания или способности к научению на основании опыта. Они были в состоянии вести лишь «растительную» жизнь. Разница между одним голубем с удаленным мозжечком и другим, с удаленной корой мозга, заключалась в том, что первая птица пыталась летать, но не могла, а вторая даже не совершала попыток.

Флоран убедительно опроверг френологию, но вопрос локализации функций мозга не был забыт, и другие физиологи вскоре обнаружили, что кора мозга имеет гораздо более высокую степень локализации функций, чем он предполагал. Например, Поль Брока показал, что относительно небольшая область левой лобной доли коры, позже названная его именем, управляет процессом речи. Более подробно о локализации функций мозга вам расскажут на занятиях по физиологии и физиологической психологии.

Бессознательно воспринимаемые сообщения и их предполагаемое влияние на человеческое поведение получили широкую известность в 1950-х гг., когда Джеймс Викари, занимающийся маркетинговыми исследованиями, заявил, что продажи попкорна и лимонада в городе Нью-Джерси сильно возросли после того как в фильм, демонстрировавшийся в городском кинотеатре, были вставлены сообщения с призывами «ешьте попкорн» и «пейте кока-колу». Предполагалось, что пока любители кино были заняты просмотром, на них неявно воздействовали сообщения, которые проходили ниже порога сознательного восприятия, но при этом были доступны для бессознательного. Несмотря на отрицательные результаты всех независимых проверок увеличения продаж и тот факт, что Викари позднее отказался от своего заявления (Pratkanis, Eskenazi, & Greenwald, 1994), защитники «бессознательно воспринимаемых записей» продолжали указывать на данный случай как на «классический» пример действенности таких сообщений. Вскоре были обнаружены и описаны в нескольких ставших бестселлерами книгах Вильсона Кея другие случаи воздействия бессознательных сообщений. Так в «Неосознаваемом соблазне» (Subliminal Seduction, 1973) Кей утверждает, что в рекламе часто используются подпороговые сообщения на сексуальную тему (например, изображение кубиков льда в форме гениталий). Идея состоит в том, что даже если люди не будут сознательно воспринимать сексуальные символы, они уловят их бессознательно. Бессознательное восприятие вызовет половое возбуждение, которое сделает рекламируемый продукт более привлекательным. Отсутствие подтверждений данного эффекта не снизило продаж книг Кея.

В 1980-х гг. в продаже появились «бессознательно воспринимаемые аудиозаписи» для самопомощи. К 1990 г. продажи достигли 50 млн долларов в год (Beyerstein, 1993). Сегодня вы сможете заказать эти записи на многих веб-сайтах или даже найти их в книжном магазине вашего колледжа. Во время прослушивания такой записи обычно слышна успокаивающая музыка или звуки природы (например, шум океана и изредка крики чаек). Неосознаваемые сообщения типа «вы можете сбросить вес» звучат неслышно, но тем не менее предполагается, что они различимы для бессознательного, которое неким неясным образом будет воздействовать на поведение человека. Росту привлекательности этих записей способствует вера покупателей в то, что минимальными усилиями можно добиться значительных успехов в самоусовершенствовании — когда сообщение укоренится в подсознании, человеком начнет управлять влиятельный внутренний голос.

Стратегия продажи «бессознательно воспринимаемых записей» требует, чтобы потенциальный покупатель был убежден в том, что данный подход имеет солидную научную основу. Действительно, проводилось множество научных исследований порогов, и некоторые из них показали, что можно воздействовать на поведение человека с помощью стимулов, лежащих ниже порога сознательного. Например, в исследовании по подпороговому восприятию участникам показывали экран с проецируемым на него словом (например, «дитя»). Слово очень быстро мигало, а его восприятие еще больше затруднялось специальными «маскирующими» стимулами. Все это делало практически невозможным распознавание и называние данного слова участниками. Таким образом, это слово лежало ниже порога сознательного восприятия и воспринималось бессознательно. Далее участникам демонстрировались быстро сменяющиеся последовательности букв. Одни последовательности

образовывали слова, другие — нет. Участники должны были как можно быстрее отреагировать на появление нужного слова. Исследователи (например, Marcel, 1983) обнаружили, что слова (например, «ребенок»), семантически связанные с бессознательно воспринятым словом («дитя»), распознавались быстрее, чем другие (к примеру, «гребенка»). «Ребенок» не распознавался быстрее «гребенки», если слово «дитя» предварительно не демонстрировалось для бессознательного восприятия. Иными словами, процесс распознавания слов «подготавливался» предварительными неосознаваемыми стимулами.

Несмотря на это, любые попытки установить связь между данным исследованием и влиянием на поведение «бессознательно воспринимаемых записей» являются необоснованными. Есть очень большая разница между влиянием экспериментатора на распознавание слов и предполагаемым воздействием на сложное поведение людей, пытающихся сбросить вес. И, как вы, наверное, догадываетесь, исследования воздействия «бессознательно воспринимаемых записей» на поведение людей показали полное отсутствие такового. Все наблюдаемые изменения возникают вследствие влияния других факторов, например ожидания человеком определенного результата. Так, в исследовании Гринвальда с соавторами (Greenwald, Spangenberg, Pratkanis, and Eskenazi, 1991) весьма остроумно показан эффект ожидания. Участникам раздали «бессознательно воспринимаемые записи» и сказали, что они улучшают память и повышают чувство собственного достоинства. Некоторым из участников выдали записи для улучшения памяти, но сказали, что они предназначены для развития чувства собственного достоинства, другим сказали, что записи должны улучшить память, тогда как на самом деле им выдали записи, развивающие чувство собственного достоинства. Вы наверняка догадываетесь, что произошло дальше. Чувство собственного достоинства увеличилось у тех, кто получил записи для развития этого чувства и был правильно проинформирован, но оно в той же степени увеличилось и у тех, кто получил записи, улучшающие память, но был в нужном направлении дезинформирован. То же случилось с теми, кто получил, и с теми, кто думал, что получил записи для развития памяти — результаты одинаково улучшились и в том и в другом случае. Как видно, значимым оказалось лишь ожидание участников. Если они ожидали усилить чувство собственного достоинства (или память), то это чувство (или память) усиливались вне зависимости от того, какие записи они в действительности слушали. Аналогично тому как исследование Флорана показало несостоятельность френологии, современные исследования продемонстрировали псевдонаучность метода «бессознательно воспринимаемых записей для самоусовершенствования».

Опора на единичные свидетельства

Вторая отличительная черта псевдонауки, помогающая понять ее популярность, — это отсутствие критики **единичных свидетельств** и опора на них при доказательстве теорий. Так, данные Галля представляют собой набор примеров, поддерживающих его теорию: вор с большой зоной «жадности», священник с чрезмерно развитой и выступающей зоной «благоговения», проститутка с непомерной зоной «влюбчивости». Защитники бессознательного восприятия используют тот же подход. Их книги наполнены отзывами людей, которые улучшили жизнь благодаря прослушиванию аудиозаписей. Например, тем кто хочет сбросить вес и знает, как давно

проводятся исследования способов быстрой потери веса, информация о том, что кто-то сбросил 20 фунтов, просто лежа в шезлонге и слушая записи крика чаек, может показаться очень заманчивой. Яркие единичные свидетельства мгновенно действуют на некритичного читателя.

Нет ничего неправильного в сборе информации, подтверждающей теорию, и даже единичные свидетельства анекдотического характера, вроде только что упомянутого, не должны отбрасываться автоматически. Проблема возникает только если целиком опираться на подобные случаи или уделять им больше внимания, чем они заслуживают, а именно так обычно и поступают с единичными свидетельствами—к ним относятся избирательно, и не укладывающиеся в схему примеры просто игнорируются (можно рассматривать это как пример искажения через подтверждение). Действительно, может быть несколько воров с особой формой черепа, но чтобы установить связь между формой черепа *X* и склонностью к воровству, необходимо узнать: а) сколько воров имеют другую форму черепа и б) сколько людей с формой черепа *X* не являются ворами. Без такой информации невозможно определить, есть ли что-нибудь особенное в том, что один или несколько воров имеют форму черепа *X*. Аналогичная проблема возникает с «бессознательно воспринимаемыми записями» — в рекламе вряд ли будут описаны случаи, когда люди покупали записи и набирали вес.

Еще одна причина не верить блестящим оценкам метода или теории состоит в том, что они часто являются следствием одного феномена, с которым хорошо знакомы социальные психологи. Он называется **оправдание усилий** (Aronson, 1999a). Теория когнитивного диссонанса Леона Фестингера говорит, что после того, как люди приложили заметные усилия для достижения цели, им необходимо убедить себя в том, что эти усилия были не напрасны. Потратив 30 долларов на записи для бессознательного восприятия и несколько десятков часов на их прослушивание, не хочется думать, что с трудом заработанные деньги были просто выброшены, а время потрачено зря. Чтобы уменьшить дискомфорт, мы убеждаем себя в том, что потратили деньги и время на стоящее дело.

Уклонение от опровержений

Из главы 3 вы узнаете, что одним из признаков хорошей научной теории является качество формулировок, достаточно точных для того, чтобы подвергнуть теорию самой строгой из возможных проверок — проверке опровержением. Особая организация теории позволяет сделать предположения и провести эксперименты, результаты которых будут подтверждать или опровергать эти предположения. В случае псевдонауки подобная проверка невозможна, хотя с первого взгляда кажется, что френологию и воздействие «записей для бессознательного восприятия» опровергнуть несложно. И действительно, научное сообщество убеждено, что данные теории уже опровергнуты.

Сторонники френологии и «бессознательно воспринимаемых записей» были вынуждены противостоять опровержениям и сопутствующему им скептицизму истинных ученых. Не у всех воров справа на черепе имеется особая выпуклость и не все слушавшие записи разбогатели. Защитники отвечают на это тем, что вместо того, чтобы признать очевидные противоречия, ставящие теорию под сомнение обходят проблемы, слегка изменяя теорию, добавляя к ней новые положения и при-

способливая ее для обнаруженных фактов. Таким образом очевидное опровержение выдается за дополнительное подтверждение теории. Например, если у известного пацифиста ярко выражены области деструктивности, опытный френолог найдет у него еще более заметные зоны осторожности, доброжелательности и почтительности, которые, по его словам, будут перекрывать очевидные тенденции к насилию. А на тот факт, что в ходе специально проведенного эксперимента не было обнаружено воздействия «бессознательно воспринимаемых записей», защитник теории ответит, что записи влияют на подсознание, а оно недоступно для традиционной науки. Далее он скажет, что записи выполняют свою функцию (только взгляните на наши отзывы), и если наука не может обнаружить факты, подтверждающие их воздействие, то это проблемы науки. Более того, если кто-то после прослушивания записей не достиг обещанного, то, вероятно, у него подсознательный блок мотивации. Таким образом, псевдонаука объясняет, или точнее обходит, все возможные факты. Но теория, которая объясняет все факты, — это не теория, так как с ее помощью невозможно сделать никаких конкретных предположений. Если у пацифистов бывают как большие, так и малые зоны деструктивности, то как можно на основании этого утверждать, что кто-то является пацифистом в силу особенностей мозга? Соответственно, если теория по определению недоступна для научного опровержения, то на основании чего тогда можно продемонстрировать ее истинность?

Другой способ, используемый псевдонаукой для уклонения от опровержения — это размытость формулировок в отчетах об исследованиях. Одной из важных особенностей науки является то, что результаты исследований общедоступны и публикуются в книгах и журналах. Еще более важно то, что ученые описывают свои исследования достаточно точно, так чтобы другие исследователи при желании могли их повторить. В псевдонауке этого не бывает, отчеты об исследованиях обычно неясны или неполны и, как было отмечено ранее, зависят от единичных свидетельств.

Сведение сложных явлений к упрощенным понятиям

В завершение отметим еще одну черту псевдонауки, заключающуюся в том, что очень сложные явления (причины поведения, факторы развития личности или факторы, вызывающие серьезные изменения в жизни) сводятся к упрощенным понятиям. Подобный подход весьма популярен, особенно если речь идет о психологии. Для человека естественно пытаться объяснять и изменять поведение, и если весь процесс можно упростить до измерения объема головы, прослушивания записей, определения астрологического знака или расшифровки почерка, то этим займутся очень многие. Обратите внимание, что в псевдонауке упрощенность идей маскируется видимой сложностью измерений. Так, френологи для выявления способностей проводили серию детально разработанных измерений черепа, а графологи, которые утверждали, что по почерку могут определить все, что вы хотите о себе узнать, замеряли десятки его характеристик (например, углы наклона).

Обобщая сказанное ранее, можно выделить следующие черты псевдонауки: а) наукообразие, б) опора на единичные свидетельства, нарушающая правила доказательства теорий, в) размытость формулировок, мешающая истинной проверке теории, и г) излишнее упрощение сложных процессов. Вследствие большого интереса людей к причинам того или иного типа поведения псевдонаучные исследования в области психологии проводились во все исторические эпохи, и, возможно, имен-

но этот интерес способствует тому, что многие люди сегодня не замечают слабости и несостоятельности псевдонаучных теорий. Развитие навыков научного мышления, к которому ведут занятия методикой исследований, позволит вам отличить истинную психологию от подделок.

Задачи исследований в психологии

В психологии перед научным исследованием стоит четыре взаимосвязанные задачи. Исследователи хотят добиться полного описания поведения, объяснения и прогнозирования его, а также с помощью сделанных ими открытий принести пользу людям.

Описание поведения

Создать **описание** в психологии означает выявить регулярные последовательности событий, включая стимулы или внешние факторы, а также ответные реакции или поведение. Например, описание агрессивного поведения у некоторых видов приматов включает набор ситуаций, в которых борьба наиболее вероятна (например, драка за еду), типы сигналов угрозы, предшествующих драке (например, оскал зубов), и виды самой драки (например, удары по зонам, не представляющим жизненной важности, — плечам и бедрам). Описание также включает классификацию (в нашем случае видов агрессивного поведения, например, борьба и хищничество). Составление ясных и точных описаний — это первый шаг в любых научных изысканиях, без него невозможно предсказать или объяснить поведение.

Прогнозирование поведения

Существование **законов** поведения означает наличие постоянных и предсказуемых взаимосвязей между переменными. В зависимости от прочности этих взаимосвязей можно осуществлять **прогнозирование** с той или иной степенью вероятности. Например, после описания нескольких сражений между приматами, становится ясно, что если двое животных подрались из-за еды и одно из них победило, то они не станут больше драться. Если оба животных одновременно заметят банан, то победитель ограничится угрожающими жестами, а побежденный, скорее всего, удалится. Если такие события наблюдается достаточно часто, то исследователи предскажут подобное поведение при будущих встречах этих животных, а также, обобщив эти результаты, предскажут поведение других пар животных, из которых одно является победителем, а другое побежденным.

Объяснение поведения

Третья задача экспериментаторов — объяснение. Объяснить поведение означает найти причины его возникновения. Понятие причинности весьма сложно, философы на протяжении многих веков размышляли о нем, и психологи-экспериментаторы хорошо осознают предположительный характер всех объяснений. Обычно исследователи пытаются выявить, что *X* является причиной *Y*, проводя эксперименты, в которых меняются параметры *X*, осуществляется управление всеми внешними факторами, способными повлиять на результат, а также наблюдается соот-

ношение вероятности появления У и вероятности случая. Таким образом, по изменению *X* можно предсказать изменение У. В этом случае говорится, что *X* и У *ковариантны* друг другу, или, иначе, появляются вместе, а так как *X* возникает раньше, то считается, что *X* является причиной У. Более того, исследователи настолько уверены в причинно-следственном объяснении явлений, что считают, что а) объяснение имеет смысл, если оно опирается на некоторую теорию или набор установленных законов, и б) все остальные возможные объяснения появления У в присутствии *X* могут быть отброшены. Процесс построения теории и разработки на ее основании эмпирических исследований, а также особенности влияния этих исследований на дальнейшее развитие теории рассматриваются в главе 3, а пока следует отметить, что процесс установления причинно-следственных связей сложен и включает в себя ковариацию, экспериментальное управление, временные последовательности с эффектом предшествующей причины, теоретическую структуру и исключение альтернативных объяснений.

Управление поведением

К задаче управления поведением иногда относятся скептически, так как нередко возникает необоснованное мнение, что психологи специально и, возможно, злонамеренно управляют жизнями людей (подробнее см. вставку 11.3 главы 11). В действительности управление означает применение законов поведения, открытых в ходе психологических исследований. Психологи предполагают, что знания, полученные с помощью научных исследований, помогут людям изменить свою жизнь к лучшему. Например, исследование факторов, вызывающих депрессию, даст врачам возможность помочь страдающим ею, а исследование агрессивности поможет родителям воспитывать детей. Эта задача редко бывает главной для исследований, но ученые о ней всегда помнят.

Страсть к исследованию (часть 1)

В начале этой главы приведены доводы в пользу того, почему занятия экспериментальной психологией необходимы студентам-психологам. Кроме того что данный курс традиционно преподается на психологических факультетах и того очевидного факта, что он является первым шагом для всех желающих стать психологами-исследователями, было сказано, что он поможет лучше понимать другие психологические дисциплины и критически воспринимать информацию о любых исследованиях, повысит шансы поступить в аспирантуру или получить работу, а также даст понимание особенностей научного мышления. Все это верно, но самая простая и наиболее важная причина занятий научным исследованием заключается в том, что исследование — это очень интересное и увлекательное занятие. Оно требует большого напряжения, иногда вызывает разочарование, а долгие часы, проведенные в лаборатории, могут быть весьма утомительными, но немногие исследователи захотят променять свою работу на другую. Что может принести большее удовлетворение, чем выдвижение гипотезы о причине поведения, последующая ее проверка путем исследований и получение совпадающих с ожиданиями результатов? Кого не взволнует мысль о возможности сделать открытие о причинах поведения, которое поможет улучшить жизнь людей?

Такое отношение к исследовательской работе как к идеальному содержанию жизни выражено в заключительном параграфе главы, написанной бихевиористом Э. К. Толменом для серии изданий о теоретических достижениях психологии. Эта глава была одной из последних его работ — она была напечатана в 1959 г., который стал годом смерти ученого. После изложения своей известной теории обучения Толмен пишет:

> Эта система (его теория) может соответствовать не всем канонам науки, но это не так важно. Мне представляется, что во всех науках, а в особенности в психологии, так много неясного и еще не открытого, что ученому лучше всего... следовать своим собственным устремлениям и своим озарениям, даже если они недостаточно верны. Я полагаю, что в действительности мы все заняты именно этим. *В конце концов, единственно истинный критерий — эта интерес. Мне было интересно.*
>
> Tolman, 1959, p. 152; курсив мой. — *Дж. Г.*

Мне бы хотелось завершить эту вступительную главу двумя небольшими примерами того, как психологи-экспериментаторы почти полностью посвящают себя работе и получают от этого глубокое удовлетворение.

Элеанор Гибсон

23 июня 1992 г. президент Джордж Буш вручил Элеанор Гибсон Национальную медаль за научные достижения *(National Medal of Science)* — наивысшую награду для ученого. Восьмидесятидвухлетняя Элеанор Гибсон была награждена за то, что посвятила всю свою жизнь исследованиям в области психологии развития и изучению всевозможных вопросов, от обучения чтению до восприятия пространственной глубины. Студентам она известна своими исследованиями «визуальной пропасти».

Элеанор Гибсон была увлеченным исследователем, не отступающим перед лицом самых серьезных препятствий — а для нее главным препятствием была дискриминация женщин-ученых. Она обнаружила это, приехав в Йельский университет в 1935 г. для работы в лаборатории по изучению приматов Роберта М. Йеркса. (Йеркс был известен своей работой в областях сравнительной психологии и тестирования умственных способностей.) Элеанор Гибсон была поражена первой встречей с ним. Как она позднее вспоминала, «он встал, подошел к двери, открыл ее и сказал: "В моей лаборатории женщин нет"» (Gibson, 1980, р. 246).

Не упав духом, Гибсон убедила крупного бихевиориста Кларка Халла в том, что она способна быть ученым, и защитила у него докторскую диссертацию. В конце 1940-х гг. она вместе со своим мужем Джеймсом Гибсоном (еще одно известное имя, на этот раз в связи с исследованиями восприятия) приехала в Корнелльский университет, в котором проработала 16 лет научным сотрудником, не получая зарплату, прежде чем ее признали профессором[1]. Именно в это время, когда ее статус был неопределен, она завершила работу по изучению развития восприятия. Радость и волнение чувствуются в ее описании первых экспериментов с «визуальной пропастью».

[1] В Корнелльском университете ей все это время не платили за работу, но она получала гранты за многочисленные удачные исследования, проведенные и описанные ею (например, гранты Фонда Рокфеллера, Национального научного фонда, Министерства образования США).



Рис. 1.3. Элеанор Гибсон получает Национальную медаль за научные достижения в 1992 г.

Ее проект состоял из серии исследований, посвященных развитию восприятия, проводимых на крысах. Работа велась совместно с ее коллегой по Корнеллу Ричардом Уолком, который, также как и она, интересовался вопросами отражения пространственных отношений в зрительном восприятии. Во время службы в армии Уолк изучил методику обучения прыжкам с парашютом, а впоследствии на корнелльской «ферме-полигоне» Гибсон наблюдала, как новорожденные козлята ведут себя, чтобы не упасть с поднимающейся платформы. Стоит отметить, что она страдала «развившейся после посещения Гранд-Каньона боязнью высоты» (Gibson, 1980, р. 258). Вместе с ассистентом они

> наскоро собрали прибор, состоящий из стеклянной платформы, поддерживаемой перекладинами. Одна половина стекла была снизу затянута бумагой, а через вторую в нескольких футах внизу был виден пол.
>
> Первую попытку совершили несколько крыс, оставшиеся после других экспериментов... Мы опустили платформу так, чтобы расстояние от пола до верхней плоскости стекла составило три дюйма, и посадили туда крыс. Нас интересовал вопрос, будут ли крысы случайным образом выбирать сторону для спуска.
>
> Случившееся было гораздо интереснее наших предположений — все крысы спустились с той стороны платформы, которая была затянута бумагой. Мы быстро поместили лист бумаги иод вторую половину платформы и повторили эксперимент. На этот раз крысы спускались с обеих сторон. После этого мы построили соответствующий прибор с управляемым освещением и т. д. *Все работало замечательно.*
>
> Gibson, 1980, р. 259; курсив мой. — *Дж. Г.*

В 1960 г. Гибсон и Уолк продолжали эксперименты с различными видами животных и людьми. Исследования «визуальной пропасти», показавшие, что уже младенцы в возрасте 8 месяцев не хотят перемещаться на «сторону обрыва», даже если там находится мама, теперь хорошо известны всем студентам, изучающим психологию.

Б.Ф. Скиннер

Если вы попросите студентов назвать имя известного психолога, кроме Фрейда, многие из них назовут Б. Ф. Скиннера (рис. 1.4) — возможно, самого известного психолога. Его работа по исследованию оперантного обусловливания положила начало целому направлению в экспериментальной психологии — «экспериментальному анализу поведения». Основные идеи этого направления рассматриваются в главе 11.

Трехтомная автобиография Скиннера освещает удивительные этапы его жизни и работы. Приведенная далее цитата демонстрирует его почти детскую завороженность открытиями в области изучения поведения. Это было написано Скиннером, после того как он получил докторскую степень в Гарварде и грант от Национального исследовательского совета и остался работать в этом университете в качестве исследователя. В начале 1932 г. он изучал обусловливание и экспериментальное торможение реакций. Скиннер писал:

> Первую кривую торможения я получил случайно. В эксперименте крыса нажимала рычаг, чтобы получить пищу, но в какой-то момент аппарат, выдающий кусочки еды, сломался. Меня не было в это время, и когда я вернулся, то *обнаружил замечательную кривую*. Крыса продолжала нажимать на рычаг, даже не получая за это пищу...
>
> Изменение поведения было более упорядоченным, чем угасание слюнного рефлекса в исследованиях Павлова, и я был ужасно взволнован. Был вечер пятницы, и в лаборатории не было никого, кому я мог бы рассказать о результатах. Все выходные я переходил улицы особенно осторожно, чтобы уберечь мое открытие и не дать ему исчезнуть в результате моей внезапной смерти.
>
> Skinner, 1979, p. 95; курсив мой. — *Дж. Г.*



Рис. 1.4. Молодой аспирант Б. Ф. Скиннер в Гарварде, ок. 1930 г.

Обратите внимание, что Скиннер и Гибсон оба используют слово «замечательный». В эксперименте Гибсон с визуальными пропастями «все работало замечательно», а Скиннер обнаружил «замечательную кривую». Речь отражает силу эмоций, которые нередко испытывают ученые-исследователи во время работы.

Скиннер был по-хорошему скептически настроен по отношению к тем, кто описывает научные методы как некий набор необходимых действий (в том числе по отношению к большинству авторов книг по методике исследований). В статье, посвященной изобретению прибора, названного его именем (проблемный ящик Скиннера), он перечислил неформальные «правила» научного проведения эксперимента (Skinner, 1956). В одном из них отразились любопытство и страсть этого человека с сильнейшим научным мышлением: «Если вы обнаружили нечто захватывающее, отбросьте все остальное и изучайте именно это» (p. 223).

Вы изучите методы и техники экспериментальной психологии и узнаете о других ученых, влюбленных в свою работу. Я очень надеюсь, что, дочитав книгу, вы почувствуете страсть к исследованиям и захотите внести свой вклад в развитие знаний о причинах человеческого поведения. Я не стремлюсь к тому, чтобы вы забросили все остальные занятия, но надеюсь, что вы найдете психологические исследования настолько увлекательными, что захотите «отбросить все остальное и изучать именно это».

Резюме

Зачем нужен этот курс?

Курс исследовательских методов — один из основных в учебном плане психологов. Его должны изучать студенты всех психологических специальностей, так как именно он дает необходимые знания о принципах психологического исследования, является основой для изучения других психологических дисциплин, помогает выработать критическое отношение к информации о законах поведения, очень важен для поступления в аспирантуру и развивает научное мышление.

Пути познания

Знание об окружающем нас мире обычно определяется имеющимся опытом и его интерпретацией, зависит от доверия к авторитетам или от других формирующих убеждения факторов, а также от способности рассуждать. Такие источники знаний весьма важны для нашей жизни, но нередко опора на них приводит к ошибкам. Опыт подвергается когнитивным искажениям (существуют, например, ригидность мнения, переоценка по запоминаемости, искажение по подтверждению), авторитеты могут быть не правы, и, хотя здравый смысл и логика равно необходимы для критического мышления, рассуждения, не подвергнутые эмпирической проверке, могут быть весьма далеки от истины. В основе познания законов поведения учеными-психологами лежит научный тип мышления.

Особенности научного типа мышления в психологии

Психологи предполагают, что человеческое поведение предсказуемо и строится по определенным законам, которые могут быть открыты с помощью научных методов. Исследователи используют метод объективного наблюдения (его результаты могут

быть подтверждены более чем одним наблюдателем). В ходе развития психологических методов стремление к объективности привело к переходу от метода интроспекции (особой формы самоотчета) к методам измерения конкретных характеристик поведения. Кроме того, выводы о причинах поведения должны опираться на получаемые данные, а поиск ответов на поставленные эмпирические вопросы должен проводиться с опорой на научные методы. Психологи-исследователи — скептические оптимисты, они относятся оптимистически к возможности открытия законов поведения и скептически к утверждениям, не имеющим достаточной эмпирической базы.

Психологическая наука и псевдонаука

Очень важно уметь различать истинную науку и псевдонауку. Последняя характеризуется наукообразием при отсылках к единичным недостоверным свидетельствам, неясным формулированием теорий, делающим невозможной их адекватную проверку с помощью научных методов, а также тенденцией к объяснению сложных явлений через излишне простые понятия. Френология XIX в. и «бессознательно воспринимаемые аудиозаписи» XX в. иллюстрируют особенности псевдонауки.

Задачи психологии

Психологические исследования нацелены на создание ясных подробных описаний типа поведения, на выявление законов, позволяющих ученым прогнозировать его с вероятностью выше вероятности случая, и на раскрытие его причин. Результаты психологического исследования должны быть применимы к непосредственному регулированию поведения — четвертой задачей является управление им.

Страсть к исследованию

Ученые-психологи очень увлечены своей работой по раскрытию законов поведения. Психологическое знание, относительно молодое, включает в себя больше вопросов, чем ответов, поэтому исследования весьма плодотворны. Удовольствие от исследования мы рассмотрели на примере жизни и работы известных психологов Элеанор Гибсон (исследования визуальной пропасти) и Б. Ф. Скиннера (создание и развитие учения о выработке инструментальных условных рефлексов).

Задания для повторения

В конце каждой главы приводятся задания для повторения. Это задания двух типов: на выбор одного ответа из нескольких и на написание коротких эссе. Прежде чем выполнять задания, внимательно изучите предшествующую им главу. В приложении Е даны ответы на вопросы и указаны номера страниц с информацией по теме данного сочинения.

Выбор ответа

1. Каков смысл тезиса о том, что научному мышлению свойственна объективность?
   1) истинные ученые никогда не позволяют априорным предпосылкам влиять на результаты работы;

2) измерения проводятся с помощью техники, поэтому воздействие человека полностью устраняется;
3) наблюдения могут быть подтверждены двумя или более наблюдателями;
4) ответы на все вопросы делаются на основании полученных данных.

2. Джон абсолютно уверен в том, что люди, получающие пособие по безработице, довольны тем, что получают «легкие деньги», и не хотят работать. Его тенденция обращать особое внимание на истории о мошенничестве социальных работников и запоминать их является примером:
   1) искажения по подтверждению;
   2) его свойства полагаться на авторитет специалистов;
   3) проблем интроспекции как метода;
   4) оправдания Джоном затраченных им усилий.
3. В соответствии с изложенным ранее психология имеет четыре **задачи. Что** из следующего соответствует категории «управление»?
   1) точная классификация различных видов шизофрении;
   2) выявление законов, предсказывающих поведение людей в различных обстоятельствах;
   3) открытие причин агрессивного поведения маленьких детей;
   4) использование результатов исследования памяти очевидцев для обучения полицейских более эффективной технике допроса свидетелей.
4. Все перечисленные ниже высказывания является характеристиками псевдонауки, кроме следующих:
   1) любые факты «объясняются» с помощью определенной теории;
   2) относительно простые явления получают очень сложное объяснение;
   3) псевдонаука выдается за нормальную научную работу;
   4) концепция опирается на единичные свидетельства.
5. К чему, по мнению Куна, приведет нежелание ученого отказаться от излюбленной теории?
   1) все обнаружат нечестность исследователя;
   2) теория не будет забыта научным сообществом до тех пор, пока не завершится ее тщательная проверка;
   3) другие исследователи заподозрят неладное и теория будет отброшена без основательной проверки;
   4) настойчивость ученого убедит остальных в истинности его взглядов.

Короткие эссе

1. В чем состоят недостатки упорства, опоры на авторитет и метода априори как способов познания?
2. Объясните, почему необходимо изучить методы исследований, прежде чем заниматься социальной психологией, когнитивной психологией или психопатологией.

3. Психологи-исследователи при работе обычно опираются на теорию статистического детерминизма. Объясните, что это означает.
4. Согласно Карнапу, без причинной обусловленности не может существовать свобода выбора. Объясните логику данного утверждения.
5. На примере техники интроспекции покажите, как психологи-исследователи понимают термин «объективность».
6. Что такое эмпирический вопрос? Приведите пример эмпирического вопроса, который был бы интересен людям, изучающим проблемы отношения религии к физическому здоровью.
7. Псевдонауку критикуют за опору на единичные свидетельства. Что такое единичные свидетельства и почему это вызывает критику?
8. Псевдонаука стремится принять наукообразный вид. Как этого добиваются при продаже «бессознательно воспринимаемых аудиозаписей»?
9. Психологические исследования имеют четыре основные задачи. Опишите их.
10. Какие условия должны выполняться, чтобы психологи-исследователи убедились в том, что открыта причина некоторого явления?

Упражнения

В дополнение к заданиям для повторения в конце каждой главы приводятся упражнения. Они представляют собой вопросы, побуждающие вас думать так, как это делают психологи-исследователи, и применять знания, полученные в ходе чтения главы. В приложении *Е* представлены ответы на некоторые из упражнений. Полные наборы ответов вы найдете у преподавателей на практических занятиях.

Упражнение 1.1. Эмпирические вопросы

Для каждого из следующих неэмпирических вопросов придумайте эмпирический вопрос на ту же тему, ведущий к потенциально интересному научному исследованию.

1. Бог умер?
2. Что такое правда?
3. Добры ли люди по своей природе?
4. Более ли нравственны женщины, чем мужчины?
5. Что такое красота?
6. Является ли разум при рождении «чистой доской» или мы рождаемся, уже имея определенный набор знаний?

Упражнение 1.2. Критическое обдумывание поговорок

Вы, вероятно, слышали поговорку «беда никогда не приходит одна». На основании своих знаний о способах познания и псевдонаучном мышлении объясните, как могло появиться подобное убеждение и почему так трудно доказать убежденному в правильности этой поговорки ее ограниченность. На основании знаний о характеристиках научного мышления укажите, что требуется сделать для проверки истинности этого утверждения.

Упражнение 1.3. Формирование строгих убеждений

Представьте себе людей, абсолютно убежденных в существовании личного Бога, направляющего их повседневную жизнь. На основании описанных в этой главе путей познания объясните, как подобное убеждение может сформироваться и чем поддерживается его существование.

Упражнение 1.4. Графология

Возможно, вы сталкивались с рекламой «науки», занимающейся изучением почерка. Нередко в такой рекламе просят прислать образец вашего почерка, в ответ на который обещают выслать описание вашей личности. Основная идея графологии заключается в предположении, что способ написания букв отражает особенности личности человека. Например, если почерк у вас мелкий или сжатый, то графологи будут утверждать, что вы стеснительны или замкнуты.

1. Охарактеризуйте основные черты псевдонауки. Покажите, как каждая из них проявляется в случае графологии.
2. Хотя планы исследований еще не обсуждались в этой книге, вы, скорее всего, уже имеете некоторое представление об эксперименте. Разработайте эксперимент для проверки вышеприведенного утверждения графологов.

ГЛАВА 2

Этика психологических исследований

Обзор задач главы

В этой главе вы познакомитесь с этическими принципами, сформулированными Американской психологической ассоциацией. Они необходимо учитывать в ходе планирования и проведения исследования, а также при составлении отчета. Этический кодекс определяет характер проведения исследований с участием как людей, так и животных[1].

Мы рассматриваем эту тему в начале книги, так как она очень важна — этические нормы непременно должны соблюдаться на всех этапах исследовательского процесса. Изучив данную главу вы сможете:

- рассказать об истории разработки этического кодекса и о различиях между общими принципами и частными стандартами;
- охарактеризовать роль Исследовательского совета *(IRB — Institutional Review Board)* и назвать требования, которые должны выполняться, чтобы *IRB* разрешил проведение исследований;
- объяснить, из чего складывается моральная ответственность исследователей. Сформулировать ее специфику в зависимости от типа испытуемых, например, взрослых или представителей определенных социальных групп;
- привести аргументы за и против использования животных в психологических исследованиях;
- объяснить, из чего складывается моральная ответственность исследователей при работе с животными;
- описать разновидности мошенничества в науке и объяснить его причины.

Этика — это система принципов, обеспечивающих моральное поведение. Вести себя этично — значит поступать правильно. Существуют этические обязательства в отношении различных областей психологических исследований. Психологи-исследователи обязаны: а) обращаться с испытуемыми людьми с уважением и не нарушать их права, б) заботиться о благополучии животных, участвующих в исследовании, и в) быть честными при сборе и обработке данных.

[1] Люди, конечно, тоже являются животными. Когда говорится об исследовании с участием животных, подразумеваются животные, не являющиеся людьми.

Прежде чем вы приступите к изучению этического кодекса *APA,* рекомендую вам прочитать вставку 2.1, в которой описано одно очень известное этически некорректное исследование. Эксперимент с маленьким Альбертом часто называют в качестве примера новаторского исследования детского страха, но он также является и примером весьма сомнительного этически поведения.

ВСТАВКА 2.1

Классические исследования - как пугали маленького Альберта

В февральском выпуске за 1920 г. «Журнала экспериментальной психологии» *(Journal of Experimental Psychology)* была напечатана статья «Условные эмоциональные реакции». Хотя методологически статья была довольно слаба (Harris, 1979), она вскоре стала одной из самых цитируемых в литературе. Ее авторами были основатель бихевиоризма Джон Б. Уотсон и Розали Рейнер - в то время аспирантка Уотсона и в будущем его вторая жена. В исследовании принимал участие лишь один испытуемый - 11-месячный мальчик, которому дали псевдоним Альберт Б. Целью исследования было проверить, можно ли создать у Альберта условную реакцию страха.

Эксперимент с маленьким Альбертом являлся частью программы изучения эмоций. До начала эксперимента было известно, что младенцы «по природе своей» почти ничего не боятся, кроме громких звуков и отсутствия опоры (Уотсон открыл это эмпирическим путем - рис. 2.1). Однако уже в раннем детстве появляется множество страхов: боязнь темноты, пауков, змей и т. д. Будучи бихевиористом, Уотсон считал, что эти страхи являются продуктами жизненного опыта, и хотел увидеть, могут ли искусственно вызванные события повлечь за собой их образование.



Рис. 2.1. Стимулы, вызывающие страх у детей

Фотографии взяты из видеозаписей, сделанных Уотсоном при исследовании младенцев (в том числе при экспериментах с маленьким Альбертом). Данные кадры опубликованы Уотсоном в его книге «Психологический уход за младенцами и детьми» *(Psychological Care of Infant and Child,* 1928). Левая фотография показывает, что произошло, когда «неподалеку от головы [ребенка] по стальному бруску ударили молотком» (p. 26). На правой фотографии показана реакция ребенка на «внезапное выдергивание из-под него простыни» (p. 27), т. е. на потерю опоры.

Сначала Уотсон и Рейнер установили, что Альберт не боится белой крысы, но сильно пугается удара молотка по стальному бруску, помещенному рядом с его головой (!). Далее исследователи попытались связать шум с крысой. Вот что произошло при первой попытке: «Белую крысу быстро достали из корзины и показали Альберту. Он попытался дотянуться до нее левой рукой. Как только он коснулся животного, за его головой ударили молотком по бруску. Младенец сильно подпрыгнул, а затем упал вперед и зарыл лицо в матрац» (Watson & Rayner, 1920, p. 4)

После нескольких повторений громкий шум стал не нужен - у Альберта выработалась прочная ассоциация, и он начал бояться крысы. Вследствие обобщения ассоциации на сходные стимулы он начал испытывать страх перед кроликом, меховым пальто и ватой.
Несомненно, невозможно призвать Уотсона и Рейнер к ответственности на основании этических норм, опубликованных через 33 года после выхода статьи об экспериментах с маленьким Альбертом. Кроме того, исследователи понимали, что найдутся люди, которые будут возражать против проведения таких экспериментов, и что «на тех, кто занимается подобными исследованиями, лежит определенная ответственность» (Watson & Rayner, 1920, p. 3). Они решили продолжать эксперименты, потому что Альберт выглядел сильным и здоровым ребенком, «в целом невозмутимым и неэмоциональным. Стабильность его состояния была одной из основных причин выбора его в качестве испытуемого... Мы чувствовали, что нанесем ему относительно небольшой вред, проводя подобные эксперименты» (p. 1-2). Уотсон и Рейнер также оправдывали свои методы тем, что Альберт приобретет эти страхи *в любом случае.* «Как только [он] поменяет безопасную детскую комнату на грубую и беспорядочную обстановку» (p. 3), он все равно приобретет их. Наука только немного раньше приведет его к этому результату.

Хотя Уотсон и Рейнер не пытались устранить страх Альберта, несколько лет спустя беседа с Уотсоном, состоявшаяся в колледже Вассар, вдохновила одну из студенток на такой эксперимент. Мэри Кавер Джонс решила показать, что приобретенный детский страх можно ликвидировать. Успешное устранение боязни кроликов у маленького мальчика (Jones, 1924) часто приводят в пример как новаторское применение техники поведенческой терапии, названной «систематическая десенсибилизация».

Читая об этических нормах *APA,* попытайтесь задать себе вопрос, возможно ли сегодня проведение такого эксперимента. Если нет, то почему? Если да, то какие меры предосторожности необходимо принять и какие изменения внести в процедуру?

Разработка этического кодекса *APA*

Психологи в США опубликовали первый этический кодекс в 1953 г. (APA, **1953).** Документ, содержащий 171 страницу, явился результатом 13 лет обсуждений, проходивших в стенах *APA,* где в конце 1930 г. был создан временный Комитет научной и профессиональной этики *(Scientific and Professional Ethics).* Комитет вскоре стал постоянным и занимался расследованиями обративших на себя внимание случаев неэтичного поведения, обычно имевших место в практике профессиональных психологов. В 1948 г. участники комитета предложили выработать официальный этический кодекс. Под руководством Никласа Хоббса был создан новый Комитет психологических этических стандартов *(Ethical Standards for Psychology),* просуществовавший 5 лет (Hobbs, 1948).

Таблица 2.1

Общие принципы этического кодекса *APA*

Принцип A. Компетентность

Психологи «осознают границы личной компетентности и пределы своих способностей» и постоянно повышают уровень своей компетентности.

Принцип B. Честность

Психологи кристально честны в «научной работе, преподавании и психологической практике» и «ведут себя по отношению к окружающим безупречно, честно и с уважением».

Принцип C. Профессиональная и научная ответственность

Психологи «придерживаются профессиональных стандартов в работе» и «несут ответственность за свое поведение».

Принцип D. Уважение к правам и достоинству людей

Психологи стараются избегать ошибок в обращении с людьми и уважают их право на «частную жизнь, конфиденциальность, самоопределение и независимость».

Принцип E. Забота о благополучии других людей

Психологи «стараются повлиять на благополучие людей, с которыми они взаимодействуют профессионально», а также не «используют и не вводят в заблуждение других людей в процессе или по завершении профессионального контакта».

Принцип F. Социальная ответственность

Психологи «применяют и делают общедоступными знания по психологии с целью повлиять на благополучие людей».

Источник: APA 1992.

Из главы 1 вы узнали, что психологи предварительно обучаются научному мышлению. Следует отметить, что члены комиссии Хоббса (в соответствии с принципом выводов на основе собранных данных) для создания кодекса провели экспериментальное исследование. Используя **метод критических случаев,** они опросили всех членов *APA* (в то время около 7500 человек) и попросили их привести примеры случаев неэтического поведения психологов, свидетелями которых были сами опрашиваемые, а также «описать, в чем, по их мнению, заключалась этическая проблема» (APA, 1953, p. vi). Было получено более 1000 ответов, в том числе несколько касающихся проведения исследований. Члены комитета разделили ответы на группы и опубликовали их в «Американском психологе» — главном журнале ассоциации *APA.* Совет директоров *APA* принял итоговую версию кодекса в 1952 г. В следующем году она была опубликована. Хотя в целом кодекс был посвящен профессиональной практике психологов, в него также вошел раздел «Этические стандарты исследований».

С тех пор кодекс был несколько раз пересмотрен, последний раз в 1992 г. Во время подготовки этой книги к печати идет еще одна переработка кодекса. Сегодня кодекс включает в себя 6 общих принципов (табл. 2.1) и 102 стандарта, разделенные на 8 основных категорий (табл. 2.2). Общие принципы имеют целью «привести психологов к высоким идеалам психологии», а стандарты — «установить *обязательные* правила поведения психологов» (APA, 1992, p. 1598; курсив в оригинале). Стандарты призваны более точно, чем принципы, определять поведение исследователей, но **я** думаю, что, изучив их, вы согласитесь с мнением социального психо-

лога Джоан Сибер (1994). Она критиковала наличный кодекс за аморфность, имеющую особенно важные следствия для тех, кто проводит психологические эксперименты без достаточного опыта (например, для вас и ваших коллег-студентов). В этой главе мы рассмотрим стандарты *APA,* касающиеся исследований с участием людей и животных[1].

Таблица 2.2

Категории этических стандартов в кодексе 1992 г.

Категория	Примеры стандартов
Общие стандарты (27)	1.02. Взаимоотношения этики и закона 1.11. Сексуальное преследование 1.27. Направления и гонорары
Подготовка, оценка или вмешательство (10)	2.03. Разработка тестов 2.05. Интерпретация результатов оценочного исследования 2.08. Получение тестовых оценок и их интерпретация
Реклама и другие публичные заявления (6)	3.01. Определение публичного заявления 3.04. Презентации для прессы 3.06. Личные ходатайства
Терапия (9)	4.02. Осведомленное согласие на проведение терапии 4.05. Сексуальная близость с пациентами или клиентами 4.09. Границы профессиональных отношений
Профессиональная тайна и конфиденциальность (11)	5.02. Соблюдение конфиденциальности 5.05. Разглашение информации 5.07. Базы данных с конфиденциальной информацией
Преподавание, руководство обучением, исследования и публикации (26)	6.03. Аккуратность и объективность в преподавании 6.06. Планирование исследований 6.23. Доверие к публикациям
Судебные разбирательства (6)	7.02. Судебные оценки 7.04. Честность и откровенность 7.06. Соблюдение законов и правил
Решение этических проблем (7)	8.02. Обнаружение этических проблем 8.05. Информирование о нарушениях этики 8.06. Взаимодействие с комитетами по этике

Примечания:

1. Число в скобках означает общее количество стандартов в этой категории.

2. Полный перечень и описание 102-х стандартов напечатан в декабрьском выпуске *American Psychologktz-A* 1992 г. (vol. 47, р. 1597-1611).

[1] Еще один полезный источник информации об этическом отношении к людям, участвующим в исследовании, — Управление корректности исследований *{Office of Research integrity)* в Министерстве здравоохранения и сферы услуг. Адрес их веб-сайта http://ori.dhhs.gov.

Этические принципы исследований с участием людей

В 1960-х гг. один из принципов первого кодекса был переработан в отдельный кодекс этики исследований с участием людей. Комитет *APA,* созданный по образцу комитета Хоббса и возглавляемый его бывшим участником Стюартом Куком, использовал тот же метод критических случаев и в 1973 г. опубликовал новый этический кодекс, специально для исследователей (APA, 1973). Он был пересмотрен в 1982 году (APA, 1982) и еще раз в 1992 г. в ходе общей ревизии. Кодекс включает набор стандартов, которые для удобства можно разделить на три основные категории — стандарты планирования исследований, стандарты проверки добровольности участия испытуемых и стандарты условий исследования.

Планирование исследований

Любое исследование человеческого поведения некоторым образом обременяет испытуемых. Как минимум участники уделяют исследованию свое время, которое могло быть потрачено иначе. Как максимум для них создают потенциально опасные ситуации. Во имя психологии **испытуемые**[1] подвергались воздействию электрическим шоком, им сообщали, что они не справились с простейшим заданием, а также приводили в замешательство различными другими способами. Что подобные процедуры могут отрицательно влиять на людей, ясно показала одна из наиболее известных в психологии серия исследований — изучение феномена подчинения, проведенное Стэнли Милгрэмом (1963, 1974).

Под видом изучения влияния наказания на обучение Милгрэм набрал добровольцев, которые должны были подчиняться требованиям экспериментатора. Участникам, выступающим в качестве учителей, приказали использовать высоковольтные электрические разряды (на самом деле никаких разрядов не было) для управления другими участниками (на самом деле профессиональными актерами), которые почти безуспешно пытались заучить последовательности пар слов. Поведение одного из участников Милгрэм описал следующим образом: «Я увидел зрелого, уравновешенного бизнесмена, уверенно и с улыбкой вошедшего в лабораторию. За двадцать минут он превратился в неуверенную заикающуюся развалину, быстро приближающуюся к нервному срыву» (p. 377). Как вы можете догадаться, исследование Милгрэма весьма спорно. Его строго критиковали за то, что он подверг добровольцев сильному стрессу, за то, что его эксперименты негативно повлияли на самооценку и чувство собственного достоинства испытуемых, и за мистификацию участников, разрушившую их веру в психологию (Baumrid, 1964).

Милгрэм столкнулся с основной дилеммой всех психологов — с необходимостью сравнивать научную ценность планируемого исследования со степенью его воздействия на испытуемых. С одной стороны, психологи-экспериментаторы твер-

[1] До недавних пор для обозначения любых участников исследования, животных или людей, чье поведение измерялось в ходе эксперимента, психологи-исследователи использовали термин **объект.** По отношению к людям сегодня этот термин не применяется, и *APA* рекомендует вместо термина «объект» использовать термин «испытуемый» или «участник исследования», так как первый не является гуманным. Однако животные могут обозначаться термином «объект».

до уверены в необходимости проведения психологических исследований самых разнообразных предметов, они убеждены в том, что отказаться от изучения вопроса — значит отречься от ответственности, возложенной на себя ученым. Если главная цель состоит в улучшении жизни людей, а для этого необходимо познание законов поведения, то очевидно, что нужно изучить их как можно более полно. С другой стороны, как мы только что видели, исследование может причинить неудобства участникам, хотя по уровню стресса, получаемого испытуемыми, очень немногие эксперименты приближаются к исследованию Милгрэма. Таким образом, при планировании исследований экспериментаторы всегда сталкиваются с требованиями, которые могут вступить в противоречие: а) получать значимые результаты, заметно увеличивающие наши знания законов поведения, и б) уважать права и благополучие участников исследования. Очевидно, Милгрэм решил, что потенциальная ценность его исследований перевешивает потенциальную опасность эксперимента. Милгрэм глубоко интересовался проблемой подчинения власти, а к исследованиям его подтолкнул вопрос о нацистском геноциде (Милгрэм был евреем). Отражает ли геноцид особенности немецкой души? Свойственна ли тенденция к подчинению власти всем нам в определенных условиях? Это важные вопросы.

Стандарты *APA* с 6.06 по 6.09 относятся к основной проблеме планирования исследований — необходимости уравновесить стремление к открытию законов поведения и потребность в защите объектов исследования. В тексте *APA* это представлено следующим образом.

6.06. Планирование исследований.

а) Психологи разрабатывают, проводят и составляют отчет об исследовании в соответствии со стандартами научной компетенции и этики исследований.

б) Психологи планируют исследования так, чтобы свести к минимуму возможность получения недостоверных результатов.

в) При планировании исследований психологи изучают их этическую приемлемость, опираясь на кодекс *APA.* Если этическая сторона неясна, психологи консультируются с исследовательскими советами, комитетами по защите животных, коллегами и другими лицами.

г) Психологи стараются обеспечить соответствующую защиту прав и благополучия участников исследования и других людей, затронутых экспериментом, а также благополучие животных, являющихся объектами исследования.

6.07. Ответственность.

а) Психологи проводят исследования в соответствии со своей компетентностью и с должным вниманием относятся к достоинству и благополучию участников.

б) Психологи отвечают за выполнение в ходе проводимых ими или под их руководством исследований этических норм.

в) Исследователям и ассистентам разрешено выполнять только те задания, для которых они специально подготовлены.

г) В ходе разработки и реализации исследовательского проекта психологи консультируются с экспертами по особым социальным группам, если эти группы изучаются или затронуты проводимым экспериментом.

6.08. Соответствие законам и стандартам.

Психологи планируют и проводят исследования в соответствии с федеральными и местными постановлениями и законами, а также профессиональными стандартами, регулирующими проведение исследований с участием людей или животных.

, **6.09. Разрешение институтов на проведение исследований.**

Психологи получают разрешение на проведение исследований у институтов или других организаций, предоставляя им подробную информацию о целях исследования. Они проводят исследование в соответствии с одобренным протоколом исследований.

Стандарт 6.06 должен гарантировать, что исследователи учтут важность этического аспекта и сделают его центральным в планировании исследований. С самого начала планирования основным вопросом должен быть баланс между научной ценностью эксперимента и правами участников. Рекомендация прибегнуть к консультации (6.06в) отражает сложность принятия решений по этому вопросу. Обычно консультация проходит в форме представления плана исследования комитету, который называется **исследовательским советом** *(Institutional Review Board)*. Он состоит не менее чем из пяти участников, обычно преподавателей с различных факультетов, хотя бы одного представителя внешней для данного учреждения организации и одного человека, не являющегося ученым (Department of Health and Human Services, 1983). Все колледжи и университеты, имеющие федеральное финансирование исследований, должны иметь свои исследовательские советы. Наибольшее распространение эти советы получили в конце 1970-х. Исследователи, обращающиеся в исследовательский совет за одобрением, обычно отвечают на серию вопросов (например, предполагает ли исследование оплату труда участников?), представляют совету подробное описание процедуры исследования, отчет о потенциальном риске для участников и о мерах по его снижению, а также типовую форму **осведомленного согласия.** Большинство советов различают «краткий обзор», подразумевающий низкий уровень риска, и «полный обзор», при котором высок уровень риска или участники эксперимента принадлежат к особым социальным или возрастным группам (например, дети или заключенные). Обычно в полных обзорах принимают участие все члены совета.

Исследовательские советы обеспечивают эффективную защиту интересов участников, исследователей и университетов, но в некоторых учебных заведениях целесообразность деятельности таких советов была поставлена под сомнение. Для этого было три причины. Первая — вопрос о степени детальности проверки исследовательских процедур и планов (Kirnmel, 1996). Исследователи законно возражают неспециалистам (например, профессорам философии), выносящим суждение о процедурах, которые они могут не понимать, к примеру о типах позиционного уравнивания (см. главу 6). С другой стороны, у исследований с плохо разработанной методикой существует своя этическая специфика. Если слаб метод, результаты

исследования могут оказаться бесполезными, а участники напрасно потратят время. Один выдающийся исследователь предложил включить в состав исследовательского совета методиста (Rosenthal, 1994). Вторая причина: некоторые исследователи убеждены в том, что очень трудно получить одобрение исследовательского совета на проведение фундаментальных исследований. Члены совета, незнакомые с отдельными областями, могут не понять важности предлагаемых экспериментов, к примеру, в фундаментальной науке, тогда как важность прикладных исследований, скорее всего, оценят по достоинству. Члену совета, читающему документ, не так просто понять необходимость исследования узнавания, при котором рассматриваются несколько изолированных переменных и все эксперименты проводятся в лаборатории с использованием искусственных материалов. С другой стороны, исследование способности узнавания с помощью записанных на видеокамеру изображений людей может показаться более важным и легче получит одобрение исследовательского совета. В-третьих, многие исследователи жалуются на то, что исследовательские советы чрезмерно беспокоятся о риске участников, а так как возможность апелляции к более высоким организациям отсутствует, совет может беспрепятственно упорствовать в своем желании блокировать исследование. Например, описанный Киммелем (Kimmel, 1996) эксперимент, в котором испытуемые должны были распознавать звуки различной громкости, не смог получить одобрения совета. Несмотря на то что громкость используемых звуков не превышала громкости разговорной речи, совет настаивал на том, что их прослушивание «связано с некоторым риском для здоровья (участников)» (p. 279). Исследователь не хотел уступать, спорил с советом в течение трех лет, а затем переключился на исследования с животными, сказав, что «состав комиссии по охране животных был более приемлемым» (p. 279). Очевидно, не все советы столь деспотичны, но отсутствие возможности апелляции действительно является проблемой.

Одна из задач этапа планирования — определение степени возможного риска для участников (стандарт 6.0бг). Иногда риск полностью отсутствует, как, например, при наблюдении поведения населения без вмешательства наблюдателя. В других случаях добровольцы могут «рисковать» или «рисковать в минимальной степени». Различие между этими двумя характеристиками не очень строгое и зависит от того, насколько создаваемые для испытуемых ситуации близки к «тем, которые обычно встречаются в повседневной жизни или в ходе принятых физических или психологических тестов и экзаменов» *(Department of Health and Human Services,* 1983, p. 279). Таким образом, участники, сталкивающиеся с ситуациями, подобными тем, которые они переживают в обычной жизни, и подразумевающими полное или почти полное отсутствие стресса, считаются «рискующими в минимальной степени». Если физический или психический риск превышает этот уровень, считается, что участники «рискуют». Например, участники исследования по спортивной психологии, посвященного тому, насколько тренировка зрительных образов влияет на атлетические способности, рискуют в минимальной степени. Но если исследование будет изучать влияние приема низких или средних доз марихуаны на вызванное тренировкой зрительных образов улучшение атлетических способностей, очевидно, степень риска будет выше.

При минимальном риске исследовательский совет обычно разрешает проведение исследований. Однако если участники «рискуют», экспериментатор должен убедить совет в том, что значимость исследования оправдывает риск, что исследование нельзя провести никаким другим способом, а также должен строго следовать принципам правильного обращения с испытуемыми и дать им необходимую информацию об исследовании.

В стандарте 6.07 говорится, что этические нормы должны выполнять все исследователи, принимающие участие в работе, а не только ее руководитель, а также что все эксперименты должны быть соответствующим образом подготовлены. Хотя в первую очередь за исследовательский проект несет ответственность руководитель исследований, все, кто связан с проведением экспериментов, должны выполнять соответствующие нормы и пройти необходимую подготовку. Для вас как для студентов, занимающихся исследованиями, этот стандарт особенно важен. Ответственность за проводимые вами в ходе занятий по методологии эксперименты в первую очередь несет ваш научный руководитель, но на вас также лежит обязанность обеспечивать защиту прав людей, участвующих в вашем проекте. До начала сбора данных необходимо также пройти подробный инструктаж (стандарт 6.07в). Для исследований, проводимых с участием «особых» популяций, предусмотрен стандарт 6.06г. Он применяется, если испытуемые по закону не имеют права самостоятельно принимать решения, как, например, дети или умственно неполноценные люди. Этот стандарт также относится к тем, кто может быть принужден стать «добровольным» участником исследования (например, к заключенным). Далее мы вкратце рассмотрим проблему согласия на участие в исследовании представителей особых социальных или возрастных групп.

Стандарт 6.08 напоминает исследователям о законах и постановлениях, относящихся к исследовательским проектам. Например, при обсуждении несколькими абзацами выше понятия «риска» участников исследования была дана ссылка на федеральные постановления Министерства здравоохранения. Другой пример таких постановлений — обязательное наличие исследовательских советов в организациях, получающих федеральное финансирование. Стандарт 6.09 имеет особое значение для исследователей, собирающих информацию за пределами научной лаборатории. Например, психологи, занимающиеся проблемами здоровья, для утверждения своей программы могут обратиться в местное медицинское учреждение: прежде чем проводить исследование, им необходимо получить его разрешение.

Проверка добровольности участия испытуемых

Второй важный для исследователей момент касается статуса участников исследования. Центральными здесь являются вопросы осведомленного согласия, мистификации участников и их права в любой момент отказаться от продолжения работы. Данные вопросы наиболее полно отражены в стандартах с 6.10 по 6.13 и 6.15.

6.10. Распределение обязанностей.

До начала исследования (кроме исследований, включающих только анонимные опросы, естественные наблюдения и т. д.) психологи заключают соглашение с участниками, в котором определяется суть исследования и обязанности обеих сторон.

6.11. Осведомленное согласие на участие в исследовании.

а) Для получения осведомленного согласия психологи используют понятный участникам исследования язык. Осведомленное согласие соответствующим образом документируется.

б) Выражаясь на понятном для участников языке, психологи информируют их о сути исследования. Они сообщают, что участники по желанию могут согласиться или отказаться от участия в исследовании, объясняют предполагаемые последствия согласия или отказа, информируют участников о важных факторах, способных повлиять на принятие решения (таких, как риск, неудобства, вредные воздействия или нарушение конфиденциальности, кроме представленных в стандарте 6.15 «Мистификации в исследованиях»), а также излагают другие аспекты, интересующие будущих участников.

6.12. Отсутствие необходимости в осведомленном согласии.

Прежде чем постановить, что планируемое исследование (например, исследования, включающие только анонимные опросы, естественные наблюдения или некоторые виды архивных исследований) не требует осведомленного согласия участников, психологи изучают соответствующие постановления и требования исследовательского совета, а также консультируются с коллегами.

6.13. Осведомленное согласие в исследованиях.

Прежде чем производить видео- или аудиозапись участников исследования, психологи получают от них осведомленное согласие. Этого не требуется только в случае исследований, основывающихся на естественном наблюдении в общественных местах и не предполагающих использование записей для опознания участников или нанесения им какого-либо ущерба.

6.15. Мистификация в исследованиях.

а) Психологи не проводят исследования с использованием мистификации, за исключением тех случаев, когда применение техники мистификации оправдано ожидаемой научной, образовательной или прикладной ценностью исследования и при этом недостаточно проведения альтернативной процедуры без мистификации.

б) Психологи никогда не скрывают от участников исследования важные аспекты, которые могут повлиять на их намерение участвовать в экспериментах, например физический риск, неудобства или неприятные эмоциональные ощущения.

в) Любая другая мистификация, являющаяся неотъемлемой частью экспериментального плана, должна быть раскрыта участникам как можно раньше, по возможности в конце их участия, но не позже окончания исследований.

В психологических исследованиях эти стандарты рассматриваются вместе, и внимание акцентируется на вопросах осведомленного согласия и мистификации. Стандарты 6.11,6.12 и 6.13 описывают принцип **осведомленного согласия,** заклю-

чающийся в том, что для принятия решения об участии в психологическом исследовании участникам необходимо предоставить полную информацию об экспериментах. Если участникам в самом начале не сообщают подробности проведения исследования или вводят их в заблуждение по поводу процедуры экспериментов, то производится мистификация (стандарт 6.15). Как же согласуются эти два противоречащих друг другу подхода?

Мистификацию предлагали заменить одним из нескольких альтернитивных вариантов, в том числе естественным наблюдением или процедурой качественного опроса. Гринберг (Greenberg, 1967), к примеру, предлагал использовать имитационную процедуру, в которой участникам предварительно сообщаются истинные цели исследования, после чего они в ролевой игре имитируют поведение ничего не подозревающих людей. Однако, проведенные исследования (например, исследование Миллера (Miller, 1972)) не подтвердили целесообразности этой идеи. Есть разница между естественным поведением и попытками вести себя соответственно ожиданиям, а поэтому неудивительно, что поведение играющих роль испытуемых и участников, ведущих себя естественно, весьма различается. Более того, участники, которым предварительно сообщили о целях эксперимента, ведут себя иначе, чем те, кто не был проинформирован. В пример можно привести исследование Гарднера (Gardner, 1978) по изучению воздействия шума как фактора стресса на различных участников — предупрежденных о шуме и не имеющих такой информации. Классические исследования показали, что шум нарушает концентрацию внимания и снижает эффективность выполнения различных задач, особенно если возникает без предупреждения. Гарднер, однако, обнаружил, что шум не оказывает неблагоприятного воздействия на тех, кто получил подробную информацию об исследовании и однозначное указание о возможности выхода из эксперимента в любой момент. Очевидно, что наличие информации усиливает у участников *чувство контроля* за ситуацией и даже непредвиденный шум не мешает им. Другие исследования (например, Sherrod, Hage, Halpern & Moore, 1977) неоднократно показывали, что усиление чувства контроля над ситуацией обычно приводит к снижению уровня стресса. Таким образом, подробно проинформированные участники исследования могут показать результаты, не отражающие реального воздействия шума. Для всестороннего изучения переменных, влияющих на взаимосвязь непредвиденного шума как фактора стресса и эффективности выполнения различных задач, необходимо прибегнуть к мистификации.

Исследование феномена подчинения Милгрэма также показывает, почему психологи нередко в начале эксперимента утаивают информацию об истинных целях исследования. Милгрэм сообщил участникам, что изучается влияние наказания на процесс обучения. Учителя (настоящие участники) сообщали ученикам набор организованных по парам слов и думали, что действительно применяют электрошок, когда ученик ошибается при воспроизведении (см. рис. 2.2). Конечно, Милгрэма интересовало совсем не обучение. Он хотел выяснить, будут ли участники а) продолжать применять электрошок по мере увеличения напряжения при контакте с теми учениками, которые явным образом чувствуют дискомфорт и не могут запомнить много, или б) не станут подчиняться экспериментатору и прекратят эксперимент. В результате лишь немногие участники отказались подчиняться приказам.

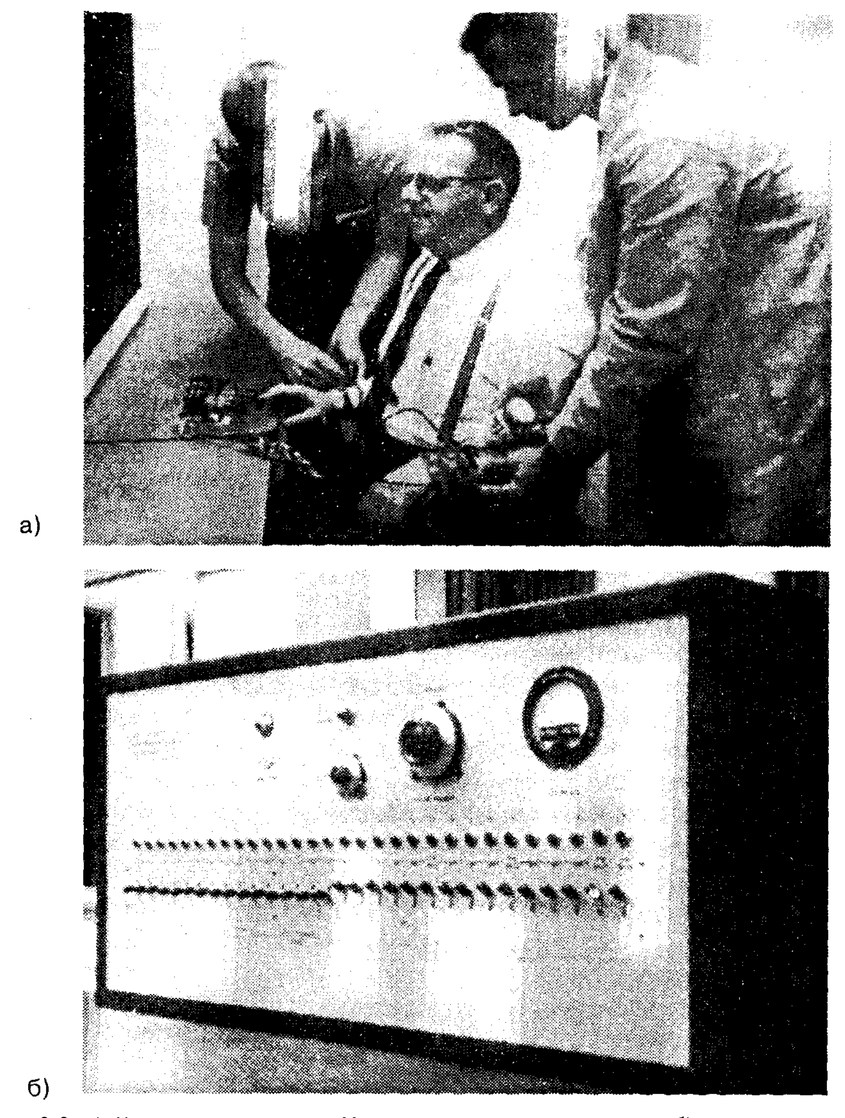


Рис. 2.2. а) «Ученика» в эксперименте Милгрэма готовят для эксперимента; б) аппарат, использованный «учителями», якобы подававший шок ученику (Milgram, 1974)

В этом исследовании 26 человек из 40 продолжали применять шок даже после того, как уровень напряжения достиг 450 вольт, и *никто* не отказался следовать приказу, пока напряжение не достигло 300 вольт (Milgram, 1963)! Если бы Милгрэм предупредил участников, что он проверяет, будут ли они подчиняться неразумным приказам, разве получил бы он такие результаты? Наверняка нет. Слепое подчинение власти никто не уважает, поэтому если участникам предварительно сообщить, что изучается подчинение, то они будут менее покладистыми. Ключевой момент состоит в том, что исследователю требуется серьезное отношение участии-

ков, их полная вовлеченность в эксперимент, а также естественное поведение. Иногда для этого просто необходимо прибегнуть к мистификации. Пожалуйста, не забывайте, что исследование Милгрэма — экстремальный пример мистификации. В большинстве психологических исследований мистификация заключается в утаивании некоторой информации об экспериментах, а не в придумывании вводящей в заблуждение истории. Это означает, что в большинстве исследований с применением мистификации при получении согласия на участие в эксперименте часть информации просто опускается, вместо того чтобы целенаправленно вводить участников в заблуждение по поводу будущих событий (Fishman, 2000).

В стандарте 6.11 говорится, что добровольцы, даже согласившись на участие в исследовании, должны знать, что они в любой момент могут безнаказанно выйти из эксперимента. Если бы исследовательские советы действовали, когда Милгрэм планировал свой эксперимент, его бы обязательно попросили изменить отдельные части процедуры исследования. Например, если учитель не решался продолжать применение электрошока (а так вели себя почти все участники), экспериментатор говорил: «Для эксперимента требуется, чтобы вы продолжали» или: «Совершенно необходимо, чтобы вы продолжали» (Milgram, 1963, p. 374). Очевидно, такое поведение экспериментатора нарушает положение кодекса о том, что участники в любое время могут выйти из эксперимента.

Вопиющее пренебрежение к осведомленному согласию при исследованиях с участием людей было продемонстрировано в Германии и на оккупированных ею территориях во время Второй мировой войны. Речь идет о медицинских экспериментах, где в качестве объектов исследования использовали заключенных концентрационных лагерей. «Во имя медицины» нацистские врачи и другие ученые (например, Йозеф Менгеле) проводили чудовищные эксперименты. Чтобы определить способности выживания человека, заключенных погружали в ледяную воду, делали им инъекции бензином или намеренно заражали смертельно опасными заболеваниями. Во время Нюрнбергского процесса эти врачи оправдывали свои действия тем, что в то время ни одно медицинское исследование по сути не проходило при добровольном согласии участников и что важность их исследований превосходит любые неблагоприятные последствия для участников. Аргументы не были приняты, ученые были признаны виновными и осуждены, а трибунал составил документ, названный Нюрнбергским кодексом. Этот кодекс стал основой для всех последующих кодексов этики медицинских исследований, а также для части этического кодекса *АРА,* касающейся согласия на участие в экспериментах и провозглашающей требования компетентности, обоснованности и добровольности согласия, а также понимания происходящего будущим участником опыта (Faden & Beauchamp, 1986).

Эксперименты, поставленные на жертвах концентрационных лагерей, конечно, представляют собой самые чудовищные примеры нарушения принципа согласия, но подобная проблема возникла также и в США. Во вставке 2.2 кратко описываются три случая из области медицинских исследований, в которых а) для изучения развития болезни дети с тяжелой степенью умственной отсталости были заражены гепатитом, б) также ради изучения протекания болезни зараженные сифилисом несколько лет не проходили лечения и были дезинформированы о состоянии своего здоровья и в) американцы получали большие дозы ЛСД, не зная об этом.

ВСТАВКА 2.2

Этика — история проблем, связанных с осведомленным согласием

Исследования врачей Третьего Рейха беспрецедентны по своей жестокости и бесчеловечности, но и в Соединенных Штатах были случаи исследований, вызвавших чрезвычайно сильную критику и сравнения, хотя и отдаленные, с нацистскими экспериментами. Три самых известных примера - это исследование гепатита в Виллоубруке, изучение сифилиса в Таскиги и проект *MK-ULTRA.*

В Виллоубруке в детском доме для детей с умственной отсталостью различной степени с 1956 по 1970 г. проводился эксперимент, в котором примерно одного из десяти вновь поступивших детей намеренно заражали гепатитом. Родители знали об эксперименте и согласились на его проведение, но позже выяснилось, что для получения согласия на них оказывалось давление. В данном эксперименте был нарушен принцип, согласно которому исследование с участием умственно отсталых может проводиться, только если оно «непосредственно связано с этиологией, патогенезом, профилактикой, диагностикой или лечением самого психического нарушения» (Beauchamp & Childress, 1979, р. 182). В Виллоубруке изучали гепатит, а не умственную отсталость.

Исследование было начато потому, что в детском доме свирепствовал гепатит, в частности, из-за большого количества детей с тяжелой умственной отсталостью, не умеющих пользоваться туалетом. В 1950 г. в учреждении постоянно проживало 5200 детей, из них 3800 имели /О ниже 20 и более 3000 не умели пользоваться туалетом (Beauchamp & Childress, 1979). При всем старании персонала в целом условия были антисанитарные, что привело к распространению заболевания. Для изучения развития болезни в контролируемых условиях исследователи преднамеренно заражали поступающих в учреждение детей и помещали их в отдельные палаты, не предоставляя лечения. Ответственные за проведение эксперимента заявляли в свою защиту, что дети наверняка заразились бы самостоятельно, так почему не сделать это специально, чтобы больше узнать о профилактике болезни? Данное исследование подверглось справедливой критике за нарушение принципа согласия, но оно действительно явилось крупным вкладом в развитие знаний о гепатите и позволило усовершенствовать методы его лечения.

Исследование в Таскиги было посвящено изучению физической деградации людей, переносящих сифилис на поздних стадиях (Jones, 1981). В начале 1930-х годов примерно у 400 нищих негров, живших в сельской местности на юге Соединенных Штатов, был обнаружен сифилис. Эти люди не получили лечения, им не сообщили о природе заболевания и его название - сказали лишь, что у них «плохая кровь». Местные врачи знали об исследовании и согласились не применять лечение. Людей настолько бедных было легко убедить периодически посещать клинику (бесплатная дорога и горячая еда) для проведения анализа крови и других исследований. Проект продолжался до начала 1970-х, хотя уже в конце 1940-х гг. было ясно, что у объектов исследования смертность в два раза выше, чем у контрольной группы, и гораздо больше медицинских осложнений (Faden & Beauchamp, 1986). Защитники исследования утверждали, что в 1930-х гг. не существовало эффективного лечения этой болезни и о ней было мало известно. Подобно экспериментам в Виллоубруке, исследование в Таскиги внесло вклад в изучение серьезного заболевания, но его значение несоизмеримо с нарушением медицинской этики.

Хотя главные исследователи и в Виллоубруке, и в Таскиги не применили адекватно принцип осведомленного согласия, они искренне хотели как можно больше узнать о двух разрушительных болезнях - гепатите и сифилисе. Третий пример нарушения принципа согласия, к сожалению, лишен оправдания пользой, принесенной развитию медицины. Проект, в котором ничего не подозревающих участников (обычно солдат) подвергали воздействию ЛСД для того, чтобы оценить возможность использования этого наркотика как оружия во время войны, был проведен Центральным разведывательным управлением. Проект запустили в начале 1950-х гг., в условиях «холодной войны» между Соединенными Штатами и Советским Союзом. Страх того, что коммунисты разрабатывают оружие, способное воздействовать на сознание, был усилен рассказами о так называемой «прочистке мозгов» американским военнопленым во время войны в Корее. В какой-то степени спровоцированное ложным сообщением разведки (о том, что Советский союз скупает мировые запасы ЛСД (Thomas, 1995)) руководство ЦРУ одобрило программу по изучению действия наркотика, в частности, возможности его использования для психического поражения врага или преодоления сопротивления пойманных шпионов. В течение 10 лет ЦРУ спонсировало многочисленные эксперименты на участниках поневоле, чаще всего солдатах, но иногда и на простых гражданах. Солдаты подписывали документ о добровольном согласии, в котором ничего не говорилось о возможном эффекте наркотика и основной целью которого было предотвратить разглашение солдатами информации о своем участии в экспериментах. Важность секретности подтверждается сделанной в ЦРУ записью, которая, в частности, гласит: «Необходимо принять меры предосторожности... чтобы скрыть данную деятельность от американской общественности... Разглашение того, что Управление предпринимает неэтичные и незаконные действия, может вызвать серьезные последствия в политических и дипломатических кругах» (Процитировано в Grose, 1994, р. 393).

Что же происходило в ходе проекта *MK-ULTRA1* Солдатам давали ЛСД, а затем изолировали их или проводили тест на детекторе лжи. Изучали также влияние многократного приема наркотика (в одном случае прием продолжался в течение 77 дней) или тайно подсыпали ЛСД мужчинам, посещающим принадлежащие ЦРУ публичные дома, где агенты наблюдали за клиентами через двусторонние зеркала (Thomas, 1995). Последний эксперимент носил кодовое название «Операция полуночный оргазм».

В ходе эксперимента два человека погибло, а на многих людей исследование оказало неблагоприятное воздействие. Ниже представлен типичный для данного проекта случай, описанный в 1994 г. в отчете Рокфеллера *(Rockfeller Report),* который был составлен по результатам расследования конгресса по делу о биологических экспериментах, в течение 50 лет финансировавшихся ЦРУ: «В 1957 г. ** добровольно согласился участвовать в особой программе по тестированию новой военной защитной одежды. За участие в исследовании ему предложили страхование жизни на значительную сумму, превосходные условия для жизни и отдыха и разрешили встречи с семьей... Затри недели тестирования новой одежды ему дали выпить два или три стакана жидкости, содержащей ЛСД. После этого мистер ** стал демонстрировать странное поведение и даже пытался покончить жизнь самоубийством. Он не знал, что в ходе эксперимента получал ЛСД, пока не услышал об этом 18 лет спустя, в 1975 г. на слушании в конгрессе» (Rockfeller Report, 1994).

«ЦРУ не потрудилось проинформировать ни конгресс, ни президента о проекте *MK-ULTRA.* Программа остановилась в 1963 г. в основном из-за того, что исследование не дало информации о возможном использовании ЛСД в военных целях. Следователи конгресса обнаружили следы проекта в середине 1970-х, составили полный отчет об этом деле и выдвинули обвинения» (Grose, 1994).

На рис. 2.3 показан пример стандартного документа о согласии на участие в исследовании. Обратите внимание на его особенности: потенциальный объект добровольно соглашается участвовать в исследовании после изучения основных целей эксперимента, процедуры и количества времени, отведенного участнику. Участникам сообщают, что полная информация об исследовании может не быть предоставлена сразу, что они в любое время могут безнаказанно оставить эксперимент, что будет соблюдаться строгая конфиденциальность и что в случае появления вопросов или сомнений по поводу исследования они смогут обратиться к консультанту. Также участникам сообщают о степени возможного риска. В завершение участники подтверждают проведение дебрифинга по окончании исследования.

Процедура

- Экспериментатор описывает общий смысл эксперимента и спрашивает о согласии участвовать в нем.
- Субъект читает документ, задает вопросы, а затем подписывает или не подписывает документ.
- После дебрифинга участник подписывает окончательную версию документа.
- Участник получает копию документа.

Эксперимент по построению когнитивных карт

Цель данного исследования - определить, насколько точно люди могут указывать географическое местоположение. Если Вы примете участие в экспрерименте, Вас попросят передвигать похожий на компас предмет так, чтобы одна стрелка указывала в направлении определенных внешних объектов. Также Вас попросят охарактеризовать степень вашей уверенности в принятом решении. Задание займет около 15 минут. Предположения, проверяемые в данном эксперименте, будут объяснены Вам по окончании участия в исследовании. Если у Вас имеются вопросы или сомнения по поводу Вашего участия или исследования в целом, обращайтесь к _____________________________

Я ознакомился с описанием эксперимента по построению когнитивных карт и добровольно соглашаюсь участвовать в нем. Я осознаю, что меня попросят указать местоположение различных географических мест и полностью объяснят цель исследования по окончании моего участия. Я осознаю, что в любой момент могу безнаказанно выйти из эксперимента и что мое участие в нем и записи об этом будут храниться строго конфеденциально.

По окончании эксперимента (обведите нужное):

требуется не требуется

краткое описание полученных результатов.

______________________________ ____________

Подпись участника Дата

****

По окончании эксперимента мне предоставили исчерпывающее объяснение исследования и ответили на все мои вопросы.

______________________________ ____________

Подпись участника Дата

Рис. 2.3. Пример документа о согласии на участие в исследовании с участием людей

Осведомленное согласие и особые популяции

Одно из последних замечаний о согласии, мистификации и возможности оставить эксперимент заключается в том, что вследствие возраста или несамостоятельности некоторые участники исследований не способны давать согласие, а на некоторых людей для получения согласия может быть оказано давление (например, на заключенных). В таких ситуациях используется особая процедура. Например, «Общество изучения развития детей» *{Society for Research in Child Development — SRCD)* следует ряду принципов, созданных на основе кодекса для взрослых. Поскольку дети могут не до конца понять документ о согласии, их родители или официальные опекуны дают согласие за них. Несмотря на это всем детям, кроме самых маленьких, исследователи обязаны сообщить об эксперименте и получить их согласие. Это значит, что исследователи предоставляют ребенку как можно больше информации, чтобы оценить его желание участвовать в эксперименте. В соответствии с кодексом *SRCD,* согласие получено, если «ребенок дал согласие на участие; при этом не обязательно полное понимание сути исследования, что совершенно необходимо для получения осведомленного согласия» (Society for Research in Child Development, 1996, p. 337). Также согласие означает, что исследователь должен наблюдать проведение экспериментов с детьми и остановить их в случае стрессового воздействия на участников. Родители могут дать осведомленное согласие на изучение влияния демонстрируемого по телевидению насилия на детскую агрессивность, но они не могут присутствовать в комнате, где показывают запись. Именно на исследователе лежит ответственность за то, чтобы освободить ребенка от выполнения задания (и устранить нанесенный вред), если уровень стресса слишком высок.

Кроме согласия кодекс *SRCD* требует, чтобы было получено согласие у тех, кто косвенно может быть вовлечен в исследование. Например у учителей, чьи ученики участвуют в экспериментах. Также в кодексе говорится о стимулах, применяемых исследователями для пробуждения желания участвовать в экспериментах, и поощрениях за выполненные задания. Поощрения «не должны сильно превышать размер обычно получаемых ребенком» (p. 337). Кроме того, исследователи не должны обещать награду, стараясь получить согласие ребенка; наоборот, о ней не следует даже упоминать до того, как родители дадут осведомленное согласие (Scott-Jones, 2000). Кодекс *SRCD* повторяет положения кодекса для взрослых, но побуждает исследователей быть еще более бдительными в отношении некоторых вопросов, например установления баланса научной значимости и риска для участников, степени оправданной мистификации, а также составления отчетов о результатах исследования.

В положениях о защите участников предусмотрены дополнения об особых слоях населения. Так, официальные опекуны должны давать действительно осведомленное согласие на исследования с участием людей, помещенных в специальные заведения (как в случае исследований в Виллоубруке). Во-вторых, необходимо убедиться, что участников не принуждают к участию в исследовании. Этой проблемы очень трудно избежать в случае тюремных заключенных, так как,

несмотря на все усилия исследователей, заключенные могут не поверить, что за отказ от участия не придется расплачиваться и что он не повлияет на будущее досрочное освобождение. В основном исследователи полагаются на простое материальное вознаграждение (например, деньги) и обещание, что участие заключенных в исследовании не будет отражено в записи о досрочном освобождении (Diener & Crandall, 1978). Аналогично требованиям кодекса *SRCD* для исследований с участием детей поощрение участников не должно превышать разумных границ. Стандарт 6.14 этического кодекса *APA* включает следующие вопросы поощрения и принуждения:

6.14. Поощрение участников исследования.

а) Предлагая профессиональные услуги в качестве стимула для привлечения участников исследования, психологи должны разъяснить суть этих услуг, а также возможный риск, обязанности и ограничения.

б) Для получения согласия на участие в исследовании психологи не должны предлагать чрезмерное или неуместное денежное или иное вознаграждение, особенно если это может привести к вынужденному участию.

Третий вопрос, касающийся особых слоев населения, — это конфиденциальность (Kimmel, 1996). Хотя обычные положения об изменении имен участников исследования остаются в силе, исследователи могут испытывать сильное желание нарушить конфиденциальность, если при этом на карту поставлены более важные вещи (например, если один заключенный сообщил о своем намерении убить другого). Как отмечено во вставке 2.2 (исследование в Виллоубруке), эксперименты с участием членов особых слоев населения должны давать информацию, приносящую пользу людям из этих слоев.

Правильное обращение с испытуемыми

Несколько разделов этического кодекса *APA* посвящено вопросам правильного и уважительного обращения с добровольцами. В них говорится о том, что по завершении исследования участники получают о нем исчерпывающую информацию, что стресс, возникший в результате экспериментов, будет снят, а их участие будет сохраняться в тайне. Ниже приведены некоторые стандарты по данной теме.

6.18. Предоставление участникам информации об исследовании.

а) Психологи обещают участникам получение необходимой информации о сущности, результатах и выводах исследования и стараются устранить любые неверные представления, возникшие у испытуемых.

б) Если научная или общественная значимость исследования оправдывает задержку или сокрытие этой информации, психологи принимают необходимые меры для снижения вероятности нанесения вреда испытуемым.

6.19. Выполнение обязательств.

Психологи принимают необходимые меры для выполнения всех обязательств, данных участникам исследования.

5.02. Соблюдение конфиденциальности.

Первейшая обязанность психологов — принимать разумные меры предосторожности для соблюдения права на конфиденциальность людей, с которыми они работают или которые у них консультируются. Они понимают, что соблюдение конфиденциальности может требоваться законом, правилами учреждения, а также профессиональными или научными отношениями.

6.17. Минимизация постороннего воздействия.

При проведении исследований психологи взаимодействуют с участниками эксперимента или другими людьми, являющимися источником данных, строго в соответствии с планом исследования и ролью психолога как ученого-экспериментатора.

Как отмечалось ранее, психолог-экспериментатор должен оценивать степень риска для участников, причем чем больше риск, тем больше усилий необходимо приложить для обоснования исследования. Проблема риска и потенциальной опасности для участников учитывается стандартом по мистификации (6.15) и стандартом 6.18, в котором говорится, что ответственность руководителя проекта не заканчивается по окончании исследования. Экспериментатор должен стараться устранить стрессовое состояние, вызванное у участников в ходе исследования, и продолжать дальнейшее наблюдение за испытуемыми, пока не убедится в их благополучии. Например, чтобы ни говорили о приемлемости экспериментов Милгрэма по изучению подчинения, ученый обращал большое внимание на эмоциональное здоровье добровольцев. По окончании исследования он послал участникам опросник с целью выяснить их переживания (84% испытуемых отметили, что они были рады участвовать в эксперименте) и пять страниц отчета о результатах исследования и их значении. Он также провел в течение следующего года дополнительное исследование, в котором 40 бывших участников наблюдались психиатром и «никаких симптомов... травматической реакции» обнаружено не было (Milgram, 1974, p. 197).

Опросы участников исследований показали, что страх ущерба, наносимого психологическими экспериментами, сильно преувеличен. Участники обычно понимают и принимают разумность мистификации (Cristensen, 1998; Fisher & Fryberg, 1994). Опрос об этической приемлемости четырех гипотетических исследований, включающих экспериментальную выработку стресса и изменение самооценки, показал, что студенты колледжа гораздо более терпимо относятся к этому вопросу, чем профессиональные психологи (Sullivan & Deiker, 1973). В табл. 2.3 показано, насколько различались ответы психологов и студентов, когда их попросили высказать свое мнение об исследовании, в котором участникам выдают тест личности и после его выполнения сообщают, что «у них довольно серьезные скрытые проблемы» (Sallivan & Deiker, 1973, p. 558). Несложно заметить, что психологи больше сомневались в приемлемости подобного исследования. Другое исследование показало, что отказ от участия в психологических экспериментах основывается скорее на предположении, что они вызовут скуку, чем на опасении, что они причинят какой-либо вред (Coulter, 1986).

Таблица 2.3

Ответы психологов и студентов об этической приемлемости исследования, призванного снизить самооценку

Вопросы	Процент психологов, ответивших «да»	Процент студентов, ответивших «да»
1. Захотели бы испытуемые участвовать в исследовании, зная его подробности?	45	80
2. Является ли мистификация неэтичной?	67	23
3. Были ли другие неэтичные моменты?	54	8
4. Было ли оправданным использование мистификации?	32	84

Как отмечалось при обсуждении осведомленного согласия, при рассмотрении проблемы мистификации важно помнить, что абсолютное большинство психологических исследований состоит из заданий, гораздо менее серьезных и с более низким уровнем мистификации, чем манипуляции самооценкой или эксперименты Милграма. Фактически в большинстве исследований с использованием мистификации участникам редко предлагают целиком выдуманные объяснения экспериментов. Мистификация обычно довольно незначительна и используется, к примеру, в заданиях на запоминание, когда субъекты заучивают и вспоминают последовательности из четырех или пяти слов, не зная, что в конце их попросят вспомнить сразу все последовательности, — это делается, чтобы не нарушить ход естественного запоминания. Во всех областях, кроме некоторых, например социальной психологии, сложная мистификация — скорее исключение, чем правило.

Стандарт 6.18 дает понятие **дебрифинга** — постэкспериментальной процедуры, в которой исследователь подробно обсуждают с испытуемым все аспекты исследования. Согласно Холмсу (Holmes, 1976a, 1976b), дебрифинг выполняет две главные функции: «демистификацию» и «десенсибилизацию». **Демистификация** означает сообщение участникам истинной цели эксперимента, а **десенсибилизация** — снижение стресса или других негативных ощущений, появившихся в ходе исследования.

Количество времени, необходимое для дебрифинга, зависит от сложности исследования, наличия и степени мистификации и уровня потенциального дистресса, ощущаемого участниками. В исследавании с мистификацией дебрифинг обычно начинается с вопроса о том, догадывались ли участники, что эксперимент имеет какие-либо другие цели, кроме заявленных в начале. Это дает возможность экспериментатору определить эффективность мистификации, а также создает необходимое вступление для сообщения истинной цели исследования. На этом этапе экспериментатор также дает обоснование мистификации (например, подчеркивая важность наблюдения естественной реакции) и начинает работу по устранению стресса, развившегося в ходе исследования. Участникам, поверившим версии, представленной экспериментатором в начале исследования, сообщается, что их поведение отражает воздействие этой версии, а не личные недостатки. Это означает, что людей, принимающих участие в самых разных видах исследований, можно заверить в том, что на их поведение сильно повлияла ситуация, в которой они на-

ходились, что их реакции не отражают индивидуальных недостатков и что другие реагировали аналогичным образом (Holmes, 1976 b). В большинстве случаев демистификация заключается в объяснении важности получения естественного поведения и обсуждении ожидаемых исследователем результатов.

В результате эффективного дебрифинга опытные экспериментаторы могут лучше понять текущие исследования и усовершенствовать будущие. Полезно выяснить мнение участников о ходе исследования — их советы по изменению процедуры позволят получить больше информации об изучаемой проблеме. Нередко рассказы испытуемых о возникших в ходе эксперимента мыслях значительно помогают интерпретировать полученные данные и планировать следующее исследование.

Правильно проводимый дебрифинг может продолжаться дольше, чем само исследование. Множество исследований показали, что участники, прошедшие подробный дебрифинг, позитивно оценивают эксперименты. В одном исследовании даже было установлено, что участники экспериментов, включающих мистификацию, оценивали образовательное значение экспериментов и собственное удовлетворение выше, чем участники, не подвергавшиеся мистификации, поскольку первые прошли более подробный и продолжительный дебрифинг (Smith & Richardson, 1983). Нельзя переоценить важность создания у субъектов хорошего впечатления об эксперименте. Участники вкладывают в наши исследования свое время, интеллектуальную и эмоциональную энергию, и за это мы находимся у них в долгу.

Также *испытуемые в некоторой степени жертвуют своей независимостью,* и мы обязаны компенсировать им эту потерю. Участникам исследования должна быть обеспечена полная конфиденциальность (см. стандарт 5.02). Это означает, что они должны быть уверены, что никто, кроме экспериментатора, не узнает их имени, а в отчетах будут фигурировать только анонимные данные. Единственное исключение составляют случаи, когда закон вынуждает исследователей открыть определенную информацию, сообщенную испытуемыми (например, жестокое обращение с детьми, положительный результат анализа на туберкулез). В исследованиях, где возможны подобные ситуации, при составлении документа о согласии экспериментаторы должны указать на ограниченность конфиденциальности (Folkman, 2000). Стандарт 6.17 касается права на частную жизнь и применяется к исследованиям, проходящим за пределами лаборатории и способным воздействовать на повседневную жизнь людей. В следующей главе вы узнаете, что, выбирая между лабораторными и полевыми исследованиями, подразумевающими вмешательство в личную жизнь людей, занимающихся обычными каждодневными делами, многие экспериментаторы предпочитают оставаться в безопасных рамках лаборатории.

Этические основы исследований с участием животных

Вы, вероятно, помните из курса общей психологии, что при проведении психологических экспериментов в качестве объекта исследования иногда используют животных. Хотя некоторые считают, что психологи больше изучают крыс, чем людей, в реальности эксперименты с животными составляют лишь малую часть всех исследований в области психологии, а точнее около 7-9% (Gallup & Suarez, 1985b).

Тем не менее такие исследования позволили сделать значительный вклад в улучшение жизни человека (Domjan & Purdy, 1995).

Есть несколько причин для использования животных в психологических исследованиях. Методологическая сторона состоит в том, что условия жизни, наследственность, рост и развитие животных несложно контролировать, а этическая в том, что, по мнению большинства психологов-экспериментаторов, при определенных мерах предосторожности животных можно использовать в таких исследованиях, которые не могут проводиться на людях. Обратимся снова к исследованию визуальной пропасти Элеанор Гибсон (Gibson & Walk, 1960). Тридцать шесть младенцев в возрасте от 6 до 14 месяцев последовательно помещали в центральную точку аппарата. Тогда как на «мелкой» стороне они демонстрировали явное желание ползать, на «глубокой» они испытывали страх и ползать не решались. Это показывает, что они были в состоянии определить высоту и сделать из этого некоторые выводы. Означает ли это, что восприятие высоты является врожденным? Нет, так как дети уже обладали опытом восприятия расстояния, полученным в течение 6-14 месяцев. Для устранения этого опыта было бы необходимо растить ребенка в условиях полной визуальной изоляции, что, конечно, невозможно. Однако эта процедура приемлема в случае животных, отчасти из-за того, что не нужно поддерживать длительную изоляцию, поскольку животные очень быстро приспосабливаются к окружающей среде, иногда в течение нескольких минут. Гибсон и Уолк экспериментировали с различными видами животных — от крыс и котят до ягнят. Они изолировали их с рождения до тех пор, пока животные не смогут самостоятельно передвигаться, а затем проверяли их реакцию на визуальную пропасть. Они обнаружили, что восприятие высоты (по крайней мере как показали эксперименты с прибором, симулирующим пропасть) у видов животных с ведущим зрительным анализатором встроено в систему зрения.

Вопрос о правах животных

Использование животных в исследованиях — вопрос спорный и эмоциональный (хотя и неновый — см. вставку 2.3). Активисты защиты прав животных осуждают эксперименты с животными (все, от медицинских до косметических). Бывали случаи разгрома лабораторий и освобождения животных. За 80-е гг., к примеру, экстремисты-экологисты разгромили 100 исследовательских центров, содержащих животных (Adler, 1992). Проблема была достаточно серьезной, чтобы вызвать изменение законодательства, и в 1992 г. вышел Закон о защите животных, специально запрещающий подобный вандализм и предусматривающий жесткое наказание за нарушение закона. Тем не менее выступления протеста, организуемые такими группами, как «Люди за этичное обращение с животными» *{People for Ethical Treatment of Animals — PETA)* или «Фронт освобождения животных» *(Animal Liberation Front — ALE)* продолжались до 90-х гг. (последняя из групп была признана ФБР террористической организацией). Известен случай убийства видной исследовательницы, занимавшейся изучением наркотиков, — номера ее домашнего и университетского телефонов стали общеизвестны, что привело к огромному количеству звонков, и на Хэллоуин ее дом был сожжен неизвестным в черной маске («Animal Rights Activity Increases», 1977).

В чем же суть возражений против использования животных как объектов исследований? Один из крайних взглядов состоит в том, что люди не имеют права считать себя выше других видов и что любые виды способны ощущать боль (Singer, 1975). При этом считается, что животные имеют такое же право на частную жизнь, независимость и неприкосновенность, как люди, и что поэтому они не должны порабощаться людьми, в том числе не должны принимать участие ни в каких видах исследований. Более умеренные группы защитников прав животных признают ценность некоторых медицинских исследований с участием животных, но отвергают возможность остальных на том основании, что исследователи причиняют бесполезную боль и страдания в тех случаях, когда альтернативное исследование может привести по сути к тем же результатам. Подобные заявления имели благотворный эффект на снижение количества бесполезных экспериментов с животными в косметической индустрии, но также сказались и на психологии. Сложилось мнение, что психологические исследования с использованием животных включают неоправданное количество повторений и направлены на тривиальные проблемы, изучение которых не приносит пользы людям. Критики исследований с использованием животных предложили, чтобы вместо изучения подопытных в лабораториях исследователи собирали информацию об их поведении, наблюдая их в естественной среде, заменяя способных к ощущениям животных на тех, кто такой способности лишены, или используя компьютерные симуляции. Каков же был ответ психологов?

Использование животных в психологических исследованиях

Большинство психологов полностью отвергают заявление, что способные к ощущениям животные имеют такие же права, как люди. Соглашаясь, что люди обязаны уважать и защищать другие виды животных, психологи считают, что люди отличаются от животных уровнем развития сознания, способностью создавать культуру и понимать историю, а в особенности способностью к моральным суждениям. Хотя животные — это удивительные создания, обладающие сложными когнитивными способностями, они «не способны быть моральными субъектами, поступать хорошо или плохо с моральной точки зрения, а также нести, выполнять или нарушать обязанности» (Feinberg, 1974, p. 46). Отличие человека от других видов животных само по себе, конечно, не обосновывает использование последних первыми, но психологи утверждают, что использование животных в исследованиях не является эксплуатацией. Более того, практический результат этих исследований приносит пользу не только людям, но *и* животным.

В психологии самым видным защитником исследований с участием животных был Нил Миллер — известный психолог-экспериментатор. Многочисленные исследования, затрагивающие самые разные вопросы, от основных процессов обусловливания и мотивации до механизмов действия биологической обратной связи, принесли ему в 1959 г. награду за выдающийся вклад в науку от ассоциации *APA* и в 1983 г. награду за выдающиеся профессиональные достижения. В своей работе «Значение исследования поведения с использованием животных» («The Value of Behavioral Research on Animals», 1985) Миллер указал, что а) активисты защиты

животных сильно переоценивают вред, наносимый животным при психологических исследованиях, б) исследования с использованием животных приносят заметную пользу для процветания людей и в) исследования с использованием животных также приносят пользу и самим животным. Говоря о проблеме причиняемого вреда, Миллер привел в пример исследование Койла и Миллера (Coile & Miller, 1984), в котором изучались последствия опубликования в журналах ассоциации *APA* 608 исследований (за 5 лет) и в ходе которого не было обнаружено *ни одного* случая протеста против жестокого обращения со стороны активистов защиты животных. Изучение таких выступлений протеста показало, что по крайней мере некоторые из мнимых случаев «жестокого обращения» никак не могли быть таковыми, а лишь казались из-за особенностей языка, использованного в описании. Например, Койл и Миллер процитировали несколько ложных заявлений, найденных в активистской литературе. Вот одно из них:«[Животных] лишают еды и воды, они страдают, а потом умирают от голода и жажды» (Coile & Miller, 1984, p. 700). Очевидно, что источником здесь явилась информация об обычных лабораторных экспериментах по формированию условных рефлексов, в которых животных лишали еды на 24 часа, после чего еду использовали как стимул для участия животных в процедуре исследования. Является ли это жестоким отношением? Вероятно, нет, принимая во внимание, что ветеринары рекомендуют кормить почти всех собак только один раз в день (Gallup & Suarez, 1985a).

Миллер указал, что во время исследований вред причиняется животным очень редко, что такие ситуации возможны, только если они оправданы приносимой экспериментами пользой, распространяющейся как на людей, так и на животных, и при этом невозможно использовать менее болезненные процедуры. Его статья от 1985 г. в основном посвящена попытке определить виды пользы, которую приносят исследования с животными. Во-первых, он отметил, что из многочисленных исследований по формированию условных рефлексов мы не только многое узнали об общих принципах обучения, но, кроме того, результаты этих исследований непосредственно применимы для решения человеческих проблем. Один из первых примеров этого — устройство, в 1938 г. разработанное и примененное Маурером и Маурером для лечения энуреза (непроизвольного мочеиспускания во время сна). Оно было разработано на основе классических работ по выработке условных рефлексов у собак Павлова. Обучающие машины и некоторые виды поведенческой терапии также основаны на принципах формирования условных рефлексов. Не так давно под влиянием исследований с животными сформировалась поведенческая медицина, использующая поведенческие принципы в традиционной терапии. Поведенческие методики, например метод биологической обратной связи, могут применяться для лечения широкого круга заболеваний, начиная с головной боли и гипертонии и заканчивая последствиями инсульта.

Также Миллер указал, что исследования с участием животных приносят пользу самим животным. Медицинские исследования значительно повышают качество ветеринарной помощи, а исследования поведения улучшают жизнь различных видов в целом. Изучение психологами поведения животных привело к улучшению условий их содержания в зоопарках, помогло выработать способы борьбы

с вредителями без применения химикатов и основанные на формировании отрицательного условного рефлекса способы отражения нападений койотов на овец, заменившие уничтожение хищников. Исследование поведения может помочь даже охране вымирающих видов. Миллер привел в пример импринтинг — особенность утят и птенцов других видов, заключающуюся в следовании за первым увиденным движущимся объектом (обычно за матерью). Изучение импринтинга привело к тому, что вылупившимся в инкубаторах птенцам кондора стали показывать куклу, напоминающую взрослого кондора, вместо того чтобы разрешать следовать за ухаживающим человеком. Это облегчило процесс адаптации к естественной среде обитания выращенным в инкубаторе птицам и значительно увеличило выживаемость птенцов видов, находящихся под угрозой вымирания.

Еще одно замечание об использовании животных в психологических исследованиях касается того, что, несмотря на выступления некоторых групп защиты животных, большинство людей считают исследования с животными весьма полезными. Например, опросы психологов (Pious, 1996 а) и студентов старших курсов психологических специальностей (Pious, 1996 b) показали, что хотя исследования, в которых животные испытывают боль и/или умирают в конце эксперимента, вызывают значительные разногласия, большинство психологов и студентов психологических факультетов считают, что психологические исследования с участием животных не только оправданы, но и необходимы. Эту точку зрения разделяют также и студенты других специальностей (Fulero & Kirkland, 1992; Gallup & Beckstead, 1988). Более того, за последние годы, несмотря на давление организаций по защите животных, в использовании животных психологами-экспериментаторами по сути ничего не изменилось. Хотя Гэллеп и Эдди в 1990 г. показали, что 14,7% учебных заведений, ранее имевших лаборатории для исследований с использованием животных, отказались от них, такой спад обусловлен не давлением защитников прав животных, а изменением исследовательских интересов и стоимостью экспериментов. Бенедикт и Столофф в 1991 г. получили аналогичные результаты в элитных колледжах. Не так давно, в 1996 г., проведенный Халлом опрос 110 заведующих кафедр вузов показал, что за 5 лет не произошло практически никаких изменений в использовании лабораторий; 47% сообщили об использовании лабораторий на момент опроса, а 50% — об их использовании пятью годами раньше. Также при опросе Халла было обнаружено, что кафедры, использующие животных для исследований, не испытывают затруднений в следовании стандандартам *APA,* и что студенты в основном положительно оценивают наличие лаборатории для экспериментов с животными. С другой стороны, Плаус (Pious, 1996a) показал, что старшие психологи активнее поддерживают исследования с животными, чем более молодые, и предположил, что в будущем количество таких психологических исследований и лабораторий для студентов психологических специальностей может сократиться[1].

[1] Более подробую информацию об этом опросе можно найти в главе 12, где он приводится как пример правильного опроса.

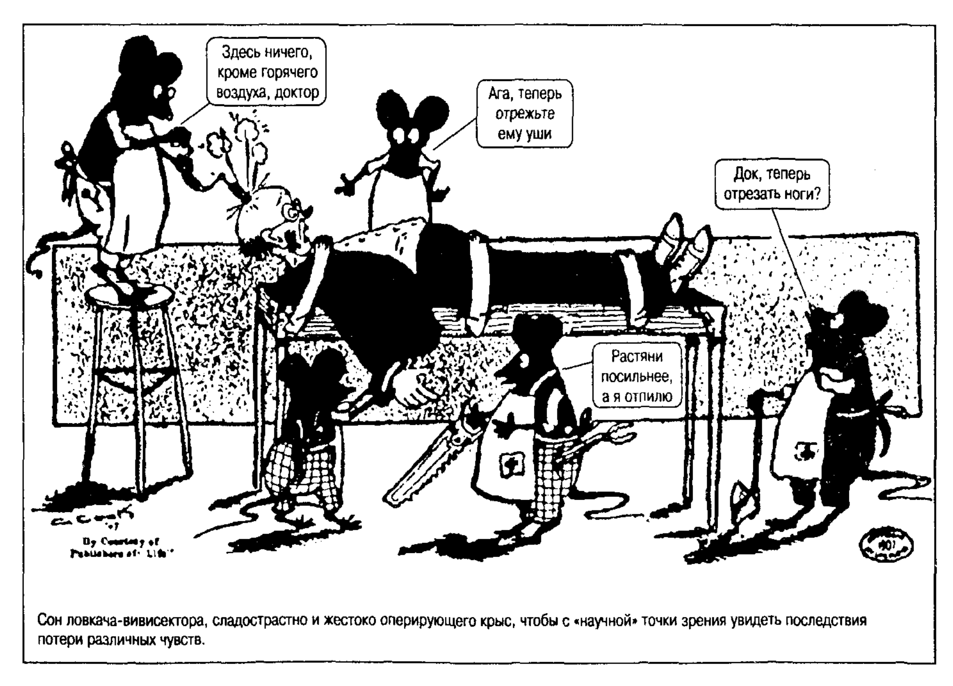
ВСТАВКА 2.3

История — антививисекция и *АРА*

Учитывая высокую активность дебатов по поводу исследований с использованием животных, вы можете подумать, что это вопрос довольно новый. Но это не так. В действительности, как это тщательно зафиксировано специалистом в области сравнительной психологии и историком Дональдом Дьюсбери (Dewsbury, 1990), вопрос имеет долгую историю.

Термин «вивисекция» происходит от латинского *vivis*- «живой» и означает проведение хирургических опереций на живых животных, обычно с определенными целями. Движение против вивисекции родилось в Англии в XIX в., где в 1879 г. усилиями активистов был принят «Акт о жестоком обращении с животными» - кодекс, сходный по духу с современным кодексом *АРА,* регламентирующим исследования с участием животных. Движение быстро распространилось и дошло до Соединенных Штатов. Американское общество против вивисекции было основано в Филадельфии в 1883 г. Противники вивисекции и исследователи, проводящие эксперименты с животными (в том числе психологи и психологи-экспериментаторы), каждый со своей стороны приводили доводы либо о чудовищных пытках, либо о научном оправдании экспериментов, которые слышны и на сегодняшних дискуссиях.

Особенно много споров вызвала серия исследований Джона Б. Уотсона (снова он). Чтобы определить, какое чувство наиболее важно при поиске выхода из лабиринта, Уотсон поставил ряд экспериментов, в которых он хирургическим путем блокировал у крыс определенные чувства, а затем изучал их поведение в лабиринте (Уотсон, 1907). 30 декабря 1906 г. исследование было напечатано в Нью-Йорк Тайме - и вызвало шумный протест. Появилась даже карикатура на Уотсона (рис. 2.4) (напечатано в Dewsbury, 1990).



Сон ловкача-вивисектора, сладострастно и жестоко оперирующего крыс, чтобы с «научной» точки зрения увидеть последствия потери различных чувств.

Рис. 2.4. Антививисекторская карикатура, изображающая Уотсона на операционном столе. Из Dewsbury (1990).

В1920 г., задолго до создания кодекса для исследований с участием людей, *APA* приняла свой первый кодекс, регулирующий исследования с использованием животных. В1924 г. был сформирован комитет, возглавленный Робертом Йерксом, и на следующий год *APA* приняла рекомендации, выработанные этим комитетом. Он предложил, чтобы лаборатории обратились к открытой политике, согласно которой «любой полномочный представитель... общественности мог бы получить разрешение на посещение лаборатории и ознакомиться с условиями обращения с животными и экспериментальными методами» (Anderson, 1926, p. 125), и чтобы журналы требовали от своих авторов четкого указания на использование в своих исследованиях гуманных процедур; чтобы психологи доказывали необходимость своих исследований, а также чтобы *APA* создала постоянный комитет по «контролю за эксперирементами с использованием животных» (Anderson, 1926, p. 125).

Кодекс *APA* для исследований с использованием животных

Стандарт 6.20 из кодекса 1992 г. дает краткое описание этических основ использования животных. Различные разделы этого стандарта были разработаны в более ранней версии документа (APA, 1985). Он касается а) необходимости обоснования экспериментов при возможности причинения вреда животным, б) вопросов правильного пути приобретения животных и обращения с ними во время и по окончании исследования и в) использования животных в образовательных целях, а не в целях исследования. Основное внимание уделено балансу между научной оправданностью конкретного проекта и требованием гуманного обращения с животными. Ниже представлены наиболее важные вопросы, описанные в стандарте.

Обоснование исследований

Аналогично тому, как исследователь, изучающий людей, должен сравнивать научную ценность исследования со степенью риска для участников, исследователь, занимающийся животными, должен доказать, что «научная цель исследования имеет достаточную потенциальную значимость, чтобы перевесить любой вред или дистресс наносимый исследуемым животным» (APA, 1985, p. 5). «Научная цель» исследования должна соответствовать одной из трех категорий: исследование должно «а) обогащать знания о процессах, лежащих в основе эволюции, индивидуального развития, устойчивости и изменчивости поведения, об управлении им или о его биологической значимости, б) увеличивать знание об исследуемых видах или в) обеспечивать результаты, могущие принести пользу для здоровья и благополучия человека и других животных» (p. 4).

Самый большой раздел документа определяет типы процедур, которые можно использовать при исследовании. В нем в основном говорится о росте требования строгого обоснования соответственно степени дискомфорта, переживаемого животными. Также указывается, что аппетентные процедуры (т. е. процедуры, использующие положительное подкрепление) должны по возможности заменить процедуры с отрицательным подкреплением, что менее стрессовые процедуры должны иметь более высокий приоритет, чем более стрессовые, и что хирургические операции требуют особого подхода и знаний. Полевые исследования должны как можно меньше нарушать покой животных, живущих в естественной среде.

Забота о животных

Специалисты, инспектирующие исследования, должны быть экспертами по данному виду животных, должны дать основательные знания всем вступающим с подопытными в контакт и знать федеральные законы об обращении с животными. Кроме того, условия содержания должны дважды в год проверяться ветеринаром, в остальное время консультирующим по телефону. Животные должны покупаться у законных поставщиков или разводиться в лаборатории. При исследовании диких животных их поимка должна осуществляться гуманным способом. Аналогом исследовательского совета по исследованиям с участием людей должен выступать «местный университетский комитет по использованию животных» (APA, 1985, p. 4). Так же как исследовательский совет, комитет по использованию животных состоит из преподавателей различных дисциплин и одного члена из внешней для университета организации. Также в состав комитета должен входить ветеринар.

По окончании эксперимента должно проводиться обсуждение альтернативы уничтожению животных. Однако иногда эвтаназия необходима «или как требование к эксперименту, или вследствие того, что она является наиболее гуманным способом избавиться от животных по завершении исследования» (APA, 1985, p. 8). В таком случае процедура должна быть «проведена гуманным способом, в соответствии с видом животных, под анестезией или способом, гарантирующим мгновенную смерть, а также в соответствии с процедурами, одобренными университетским комитетом по использованию и обращению с животными» (p. 8).

Использование животных в образовательных целях

Этические основы разработаны в первую очередь в помощь исследователям, проводящим эксперименты с животными, но животные также часто используются в образовательных целях: для демонстрации особого типа поведения, для обучения студентов процедурам исследований, а также для того, чтобы студенты могли получить непосредственный опыт изучения таких известных феноменов, как классическое и оперантное обусловливание. В противоположность ситуации исследования использование животных при обучении не приводит непосредственно к формированию новых знаний, поэтому преподавателю для достижения заданной цели необходимо использовать как можно меньше животных, а также обдумать возможность альтернативных процедур. Например, вместо того чтобы каждый семестр при чтении вводного психологического курса для демонстрации одной и той же закономерности (например, формирования условных рефлексов путем последовательного приближения к конечной цели) использовать новую крысу, преподаватель может проделать это один раз, а в будущем демонстрировать видеозапись.

Иногда живую процедуру можно заменить компьютерной симуляцией различных феноменов — существует несколько превосходных симуляций как классического, так и оперантного обусловливания. Симуляции могут быть весьма эффективны (и необходимы в случае небольших учебных заведений, которые не могут обеспечить соблюдение федеральных правил по подобающему обращению с животными), но научить компьютерную крысу нажимать рычаг — не совсем то же, что научить реальную крысу. Студенты часто переживают глубокое понимание мощи ситуатив-

ного подкрепления, когда наблюдают его лично. Непосредственный опыт работы с животными в студенческой учебной лаборатории подтолкнул многих студентов к выбору профессии психолога-исследователя (Moses, 1991).

В целом, большинство психологов одобряют использование животных для исследования поведения, осознавая при этом необходимость тщательного анализа обоснованности каждого эксперимента с животными. Исследования с участием животных внесли существенный вклад в наше понимание поведения и обещают помочь в поисках решений проблем СПИДа, болезни Альцгеймера, психических заболеваний и других проблем человечества.

Научное мошенничество

В последние годы часто шли дискуссии по поводу мошенничества в науке, но особенно пылкие дебаты вызвал вопрос, является ли оно просто случайным «гнилым яблоком» или «верхушкой айсберга», основание которого не предвещает ничего хорошего. Ясно, что ученые вообще и психологи-исследователи в частности должны быть кристально честными в своей научной деятельности. Принцип Б общего кодекса 1992 г. недвусмысленно указывает, что психологи «должны проявлять честность в науке, преподавании и психологической практике» (АРА, 1992). Более того, несколько особых стандартов кодекса 1992 г. прямо касаются мошенничества в исследованиях. Данный раздел посвящен следующим вопросам: что такое научное мошенничество? Насколько оно распространено? Почему оно возникает?

Толковый словарь *«American Heritage Dictionary»* (1971) определяет мошенничество как «преднамеренный обман, практикуемый с целью получения незаслуженной или незаконной выгоды» (р. 523). В науке распространены два основных вида мошенничества: 1) **плагиат** — преднамеренное присваивание чужих идей **и** выдача их за собственные и 2) **фальсификация данных. В** кодексе 1992 г. плагиат специально осуждается стандартом 6.22, а фальсификация данных — стандартом **6.2**1 (табл. 2.4). Проблема плагиата свойственна всем областям человеческой деятельности, а фальсификация данных встречается только в науке, поэтому следующий раздел будет посвящен именно этому вопросу.

Таблица 2.4

Фальсификация данных и плагиат: стандарты *АРА*

Стандарт 6.21. Отчет о результатах

а) Психологи не фабрикуют данные и не фальсифицируют результаты исследований в своих публикациях.

б) Если психологи обнаружат важные ошибки в опубликованных ими данных, они стараются исправить эти ошибки путем коррекции, опровержения, исправления опечаток или другими подходящими средствами.

Стандарт 6.22. Плагиат

Психологи не выдают значительные части чужих работ за свои собственные, даже при наличии ссылок на эти работы или источники данных.

Источник: АРА (1992).

Фальсификация данных

Если и лежит на науке моральный грех, то это грех отсутствия кристальной честности в обращении с данными, а отношение к данным ложится в фундамент целого здания науки. Но если подводит фундамент, все остальное рушится, поэтому безупречность данных представляет собой вопрос первостепенной важности. Данный вид мошенничества может принимать различные формы. Первая и самая экстремальная форма — это когда ученый вообще не собирает данные, а просто фабрикует их. Вторая — сокрытие или изменение части данных для лучшего представления конечного результата. Третья — сбор некоторого количества данных и додумывание недостающей информации до полного набора. Четвертая — сокрытие всего исследования, если результаты не соответствуют ожидаемым. В каждом из этих случаев обман является преднамеренным и ученые, по-видимому, «получают незаслуженную или незаконную выгоду» (т. е. публикацию).

Традиционная точка зрения состоит в том, что мошенничество встречается редко и легко раскрывается, так как фальшивые результаты не могут быть воспроизведены (Hilgartner, 1990). Это означает, что результат, полученный ученым исходя из фальсифицированных данных, не отображает эмпирической правды. Другие ученые, заинтригованные или удивленные новым открытием, попытаются, но безуспешно воспроизвести его в собственных лабораториях. Таким образом, подложное открытие будет обнаружено и отвергнуто. Подозрение о фальсификации данных появляется, когда исследователь, обнаружив нечно странное, просит показать необработанные данные, собранные в ходе экспериментов. Ученые-психологи и представители других областей науки с давних пор охотно делятся собранной информацией, и отказ показать данные создаст подозрения насчет «нового» открытия. Что от ученого совершенно естественно ожидать согласия поделиться данными, видно из этического кодекса *APA*.

Стандарт 6.25.

> После опубликования результатов исследования психологи не должны скрывать данные, лежащие в основе сделанных выводов, от других ученых, желающих проанализировать их с целью проверки выдвинутого утверждения и намеревающихся использовать данные только для этого, при условии, что возможно защитить конфиденциальность участников и если законные права на патентованные данные не препятствуют их публикации.

Кроме провала попыток повторить сделанные открытия мошенничество может **быть** обнаружено (или хотя бы заподозрено) в ходе стандартной проверки. Когда статья об исследовании представлена на рассмотрение в журнал или в агентство подано заявление на получение гранта, несколько экспертов осуществляют ее проверку, помогающую решить, будет ли опубликована статья или выдан грант. Моменты, выглядящие странно, возможно, обратят на себя внимание хотя бы одного из исследователей. Третья возможность обнаружить мошенничество — когда проблему заподозрят работающие вместе с исследователем сотрудники. Так случилось в 1980 г. в ходе одного печально известного исследования. В серии экспериментов, которые, казалось, совершили прорыв в лечении гиперактивности у детей с задержкой развития, Стивен Брюнинг получил данные, говорящие, что в этом случае

стимулирующие лекарства могут быть более эффективными, чем антипсихотические (Holden, 1987). Однако один из его коллег подозревал, что данные подделаны. Подозрение было подтверждено после трех лет расследований Национальным институтом психического здоровья *(National Institute of Mental Health — NIMH),* который финансировал некоторые исследования Брюнинга. В суде Брюнинг признал себя виновным в двух случаях представления в *NIMH* фальсифицированных данных; в ответ *NIMH* снял обвинения в лжесвидетельстве во время расследования (Byrne, 1988).

Одна из сильных сторон науки — это самокоррекция вследствие повторения экспериментов, пристальных проверок и честности коллег. И действительно, такая ее организация много раз позволяла обнаружить мошенничество, как, например, в случае с Брюнингом. Но что если эксперты не смогут обнаружить никаких следов фальсификации или если фальсифицированные результаты будут соответствовать другим, настоящим открытиям (т. е. если их можно повторить)? Если поддельные результаты согласуются с истинными открытиями, то не возникает повода для их проверки и мошенничество может оставаться нераскрытым многие годы. Вероятно, нечто подобное произошло в самом известном в психологии случае подозреваемого мошенничества («подозреваемого», так как окончательное решение все еще не вынесено).

Случай касается одного из самых известных британских психологов — Сирила Берта (1883-1971), главного участника дебатов о природе интеллекта. Его исследования близнецов часто приводят в пример как доказательство того, что интеллект преимущественно наследуется от одного родителя. Один из результатов Берта показал, что однояйцевые близнецы имеют практически одинаковые показатели *IQ,* даже если сразу после рождения их усыновили разные родители и они воспитывались в разных условиях. Многие годы никто не ставил под вопрос его данные, и они вошли в литературу, посвященную наследуемости интеллекта. Однако внимательные читатели со временем заметили, что, описывая в разных изданиях результаты, полученные при исследовании разного количества близнецов, Берт указывал *абсолютно* те же статистические результаты (тот же коэффициент корреляции). С математической точки зрения получение таких результатов очень маловероятно. Противники обвинили его в подделке результатов с целью подкрепить убеждения Берта в наследуемости интеллекта, а защитники возразили, что он собрал действительные данные, но с годами стал забывчив и невнимателен в своих отчетах. В защиту ученого также было сказано, что если бы он занимался мошенничеством, то наверняка постарался бы это скрыть (например, позаботился бы о несовпадении корреляций). Нет сомнений, что в данных Берта есть нечто странное, и даже его защитники признают, что многие из них не имеют научной ценности, но вопрос о том, было ли совершено намеренное мошенничество или все дело в невнимательности и/или небрежности, может быть никогда не решен, в частности потому, что после смерти Берта его экономка уничтожила несколько ящиков, содержавших различные документы (Kohn, 1986).

Стало очень популярно разбирать дело Берта (Green, 1992; Samelson, 1992), но для нас важно, что неправильность в данных, вызванная ошибками, невнимательностью или намеренным искажением, может остаться незамеченной, если

данные хорошо согласуются с другими открытиями (т. е. если они кем-либо воспроизведены). Именно так обстояло дело у Берта — его данные были весьма похожи на полученные в других исследованиях близнецов (например, Bouchard & McGue, 1981).

Следует отметить, что некоторые комментаторы (например, Hilgartner, 1990) считают, что кроме случая, когда фальсифицированные данные повторяют «правильные» данные, есть еще два типа причин, по которым фальсификацию могут не обнаружить. Во-первых, большое количество исследований, публикующихся сегодня, позволяет проскользнуть подложной информации незамеченной, особенно если в ней не сообщается о крупных открытиях, привлекающих широкое внимание. Во-вторых, система вознаграждений устроена таким образом, что новые открытия оплачиваются, тогда как работа ученых, занимающихся «простым» воспроизведением чужих результатов, не считается вполне творческой и такие ученые не получают академических наград. Вследствие этого некоторые сомнительные исследования могут не подвергнуться воспроизведению.

Также считается, что система вознаграждений является в каком-то смысле причиной появления мошенничества. Это мнение подводит нас к заключительному и основному вопросу — почему возникает мошенничество? Есть различные объяснения — от индивидуальных (слабость характера) до социальных (отражение общего морального упадка конца XX в.). Возложение ответственности на академическую систему вознаграждений помещается где-то в середине списка причин. Ученые, публикующие свои исследования, продвигаются по службе, получают постоянные должности, выигрывают гранты и имеют возможность воздействовать на аудиторию. Иногда постоянное «умри, но опубликуй» так сильно действует на исследователя, что приводит его (или его ассистента) к мысли нарушить правила. Это может происходить сначала в ограниченных пределах (добавление небольшого количества информации для получения желаемых результатов), но со временем процесс будет разрастаться.

Что это означает для вас — студентов, занимающихся исследованиями? По меньшей мере это значит, что вам необходимо быть добросовестными по отношению к данным, скрупулезно следовать процедуре исследования и *никогда* не поддаваться искушению фальсифицировать даже незначительный объем информации; также никогда не отбрасывать данные, полученные от участников исследования, кроме случаев, когда для этого существуют четкие указания, определенные до начала эксперимента (например, когда участники не следуют инструкциям или исследователь неправильно руководит ходом эксперимента). Кроме того, необходимо сохранять исходные данные или по крайней мере иметь их краткое описание. Лучшая защита против обвинений в том, что ваши результаты выглядят странно, — это ваша способность предоставить данные по требованию.

Важность этической основы проводимых исследований нельзя переоценить, поэтому данная глава помещена в самом начале книги. Но обсуждение этических норм не ограничивается пределами одной главы — в дальнейшем вы еще не раз столкнетесь с этой темой. Если вы, к примеру, обратите внимание на содержание, то увидите, что каждая последующая глава содержит вставку об этике, посвящен-

ную таким вопросам, как нераспространение информации об участниках полевых исследований, отбор участников, ответственное использование опросов и этическая компетентность экспериментаторов. В следующей главе, однако, мы рассмотрим проблему из другого круга — выработка идейной основы исследовательских проектов.

Резюме

Разработка этического кодекса *APA*

В соответствии с правилом психологов — полагаться на информационный подход — ассоциация *APA* разработала свой первый этический кодекс эмпирически, с помощью метода критических случаев. Кодекс для исследований с участием людей впервые был опубликован в 1953 г. и с тех пор периодически пересматривался, в последний раз в 1992 г. Кодекс состоит из общих принципов, определяющих поведение психологов (например, принцип заботы о благополучии окружающих), и особых стандартов поведения (например, стандарт обеспечения конфиденциальности), нарушение которых строго осуждается.

Этические основы исследований с участием людей

Кодекс *APA* для исследований с участием людей представляет собой руководство по планированию и проведению экспериментов. Планирование включает анализ соотношения затраты/польза, при котором сравнивается уровень риска участников и научная ценность исследования. Также кодекс требует, чтобы будущие участники исследования получили информацию, достаточную для принятия решения об участии (т. е. для осведомленного согласия). Особое внимание необходимо уделить детям и людям, которых могут вынудить на участие в исследовании (например, заключенным). Следует сообщить участникам, что они могут безнаказанно оставить эксперимент, и заверить их в полной конфиденциальности. По окончании экспериментов испытуемые должны пройти полный дебрифинг. Исследовательские советы несут ответственность за то, чтобы эксперименты с участием людей проводились в соответствии с этическим кодексом. В психологических исследованиях возможны некоторые виды мистификации, но для получения разрешения на ее применение исследователь должен убедить исследовательский совет в том, что настоящая цель исследования иначе не может быть достигнута.

Этические основы исследований с использованием животных

Разработанные *APA* основы исследований с использованием животных включают требования заботы и гуманного обращения с подопытными, являются руководством по выбору подходящей экспериментальной процедуры и касаются использования животных как для исследований, так и в образовательных целях. Хотя защитники прав животных утверждают, что психологические исследования с участием животных недопустимы, большинство психологов-исследователей считают, что такие исследования приносят пользу не только людям, но и животным.

Научное мошенничество

Плагиат — представление чужих идей как своих собственных — и фальсификация данных — подделка или изменение данных — это наиболее серьезные виды научного мошенничества. Хотя фальсификация данных часто обнаруживается при безуспешных попытках воспроизвести подложные открытия, она может остаться невыявленной, когда а) фальсифицированные открытия согласуются с истинными или б) они теряются в огромном количестве публикаций. Иногда академическая система вознаграждений становится фактором, приводящим к научному мошенничеству.

Задания для повторения

Выбор ответа

1. У Джона Уотсона и Розали Рейнер сегодня возникли бы затруднения при попытке получить от исследовательского совета разрешение на проведение исследования с маленьким Альбертом. Почему это так?
   1) изучение эмоций у 11-месячных детей неприемлемо, так как они не могут дать согласия на участие в эксперименте;
   2) ученые предложили потенциально вредную процедуру, не обеспечив устранения ущерба по окончании эксперимента;
   3) у исследования был только один объект, а следовательно, возможность обобщения результатов была ограничена;
   4) исследование подразумевало нарушение конфиденциальности, так как Уотсон и Рейнер открыли имя ребенка.
2. Для получения осведомленного согласия участникам сообщают:
   1) что эксперимент будет проводиться без мистификации;
   2) что они могут решать, принимать ли участие в эксперименте, но если они согласятся, то должны будут довести его до конца;
   3) будет ли применяться в эксперименте электрический шок;
   4) общее содержание экспериментов в случае исследования без мистификации, в противном случае предварительно *ничего* не сообщают.
3. Изучение использования мистификации в психологических исследованиях показало, что:
   1) оно не является необходимым — участники, в ролевой игре имитирующие необходимое поведение, показывают те же результаты, что и обычные добровольцы;
   2) студенты при анкетировании называют эксперименты с мистификацией неэтичными чаще, чем психологи-исследователи;
   3) оно действительно вредно. По окончании экспериментов по изучению подчинения большинству из испытуемых-объектов исследования Милграма потребовалась психологическая помощь;

4) поведение участников, подробно проинформированных о целях исследования, часто сильно отличается от поведения тех, кто не получил полной информации.

4. Все следующие пункты входят в последний кодекс *APA* по использованию животных в исследованиях, кроме:
   1) психологам больше не разрешается использовать животных исключительно в образовательных целях;
   2) можно причинять боль животным, если ценность исследования перевешивает степень риска объектов исследования;
   3) для оценки исследований с использованием животных должна существовать группа, аналогичная исследовательскому совету, который рассматривает заявки на проведение исследований с участием людей;
   4) если по окончании исследования или как часть экспериментального плана необходимо провести эвтаназию, она должна быть безболезненной.

5. Что из нижеперечисленного верно в отношении фальсификации исследовательских данных?
   1) Мошенничество в конце концов будет обязательно раскрыто, так как результаты подложных экспериментов не соответствуют полученным в истинном исследовании;
   2) она определяется как подделка всех или большинства данных исследования. Изменение малого количества данных не считается мошенничеством;
   3) она может остаться нераскрытой, если сфабрикованные результаты похожи на истинные, полученные в других исследованиях;
   4) это довольно редкое явление, потому что организация академической системы не позволяет получить вознаграждение за сфабрикованные данные.

Короткие эссе

1. Покажите отличие «общих принципов» и «стандартов» этического кодекса *APA.* Приведите примеры.
2. Опишите роль исследовательского совета и приведите две причины, по которым целесообразность его деятельности ставят под сомнение.
3. Покажите отличие согласия, которые дают взрослые, от согласия детей и объясните, насколько оба понятия важны для исследований с участием детей.
4. Перечислите наиболее важные части документа для получения осведомленного согласия, используемого при исследованиях с участием взрослых.
5. Почему в психологических исследованиях иногда используют мистификацию? Как использование мистификации согласуется с идеей осведомленного согласия?
6. Опишите две главные цели дебрифинга.
7. Какие этические принципы были нарушены в а) исследовании в Виллоубруке, б) исследовании в Таскиги и в) в проекте *MK-ULTRA?*

8. На примере исследования визуальной пропасти по Гибсон объясните, почему психологи иногда используют животных в качестве объектов исследования.
9. Приведите доводы за и против использования животных в психологических исследованиях.
10. Каковы главные особенности кодекса *APA* для исследований с использованием животных?
11. Опишите способы, с помощью которых обычно раскрывается фальсификация данных.
12. На примере случая Берта покажите, почему фальсификация данных может быть не замечена. Каковы еще две причины, по которым такой тип мошенничества может остаться нераскрытым?

Упражнения

Упражнение 2.1. Научное мышление и мистификация

Встаньте на точку зрения психолога-исследователя, способного к научному мышлению, и опишите разработку исследования для проверки приведенных ниже высказываний. Такие мнения нередко высказываются по поводу мистификации. Какие эмпирические данные вам понадобятся для проверки истинности этих утверждений?

1. Мистификация не должна использоваться в психологических исследованиях, так как если людей обманут в ходе исследования, они больше не поверят ни одному психологу.
2. Мистификацию можно заменить на игровую ситуацию, в которой испытуемым предлагают «вообразить», что они проходят эксперимент с мистификацией, и просят вести себя так, как, по их мнению, это должен делать обычный человек.
3. Психологи сами себя обманывают — большинство испытуемых догадываются об истинных целях исследования.

Упражнение 2.2. Обнаружение этических проблем

Рассмотрите каждое из следующих кратких описаний исследований по социальной психологии, имевших место в реальности. Какие стандарты могут вызвать возражения исследовательского совета (с точки зрения этического кодекса *APA)?* Укажите, какие доводы можно привести в защиту каждого из исследований перед исследовательским советом.

1. В общественном мужском туалете было проведено исследование влияния присутствия окружающих на переживаемый стресс. Один из членов исследовательской команды входил в туалет вместе с другим человеком и занимал смежный писсуар или один из более отдаленных. Объект наблюдения не знал о своем участии в исследовании. Было сделано предположение, что усилившийся стресс будет воздействовать на мочеиспускание. Спрятавшийся в одной из кабин исследователь записывал время до начала мочеиспускания и его продолжительность. Как и было предположено, мочеиспускание

было более прерывистым, когда исследователь занимал смежный писсуар (Middlemist, Knowles, & Matter, 1976).

2. В полевом эксперименте женщина — участник исследовательской команды стояла возле машины на краю дороги. У машины была спущена шина. Целью исследования было определить влияние примера на поведение проезжающих автомобилистов. Для этого за четверть мили до того места, где женщина ожидала помощи, была разыграна сцена: остановившийся водитель помогал другой женщине со спущенной шиной. Как и ожидалось, автомобилисты проявляли большую готовность остановиться и помочь, когда они только что были свидетелями чужой помощи (Bryan & Test, 1967).

Упражнение 2.3. Изучение подчинения авторитету

Представьте, что вы хотите в наше время изучить особенности подчинения авторитету и планируете использовать процедуру, подобную разработанной Милграмом. В частности, вашей целью является проверить, повторятся ли результаты, полученные в оригинальном исследовании, 25 лет спустя. Как вы обоснуете эксперимент для исследовательского комитета и какие изменения внесете в процедуру Милграма? Используя шаблон, приведенный в данной главе, разработайте документ для получения согласия на участие в вашем исследовании.

Упражнение 2.4. Стремление избежать плагиата

Студент писал работу, посвященную лжи, и читал книгу Пола Экмана «Ложь» (Paul Ekman, Telling Lies, 1985). Он прочитал о распространенной ситуации, когда человек в действительности говорит правду, но при этом так волнуется, что выглядит, будто лжет:

> ...Другой столь же важный источник проблем, приводящий к неверию в правду, — это ошибка Отелло. Она возникает, когда, пытаясь уличить во лжи, не осознают, что если говорящий правду человек испытывает стресс, то может казаться, что он лжет.
>
> ...Честные люди могут опасаться, что им не поверят, и этот страх можно принять за боязнь лжеца быть раскрытым...
>
> Я назвал эту ошибку именем Отелло, так как сцена убийства в шекспировской пьесе — это превосходный и очень известный пример подобной ошибки... (Ekman, 1985, p. 169-170).

Студенческая курсовая работа о лжи содержала следующий параграф:

> Неверие в правду — это еще одна проблема правильного распознания лжи. Иногда человек, старающийся уличить кого-то во лжи не учитывает, что если говорящий правду испытывает стресс, то может показаться, что он лжет. Этот феномен известен как ошибка Отелло. Таким образом, иногда можно говорить правду, но при этом опасаться, что вам никто не поверит. Вследствие этого вы будете выглядеть как человек, боящийся быть раскрытым.

Получив проверенную работу, студент был шокирован, что профессор обвинил его в плагиате и поставил неудовлетворительную оценку. Как вы думаете, почему это произошло? Как бы вы изменили этот параграф, чтобы обвинение в плагиате не было сделано?

ГЛАВА 3

Разработка идей для психологических исследований

Обзор задач главы

Все исследования начинаются с правильной постановки вопроса, и данная глава призвана помочь вам научиться делать это. Глава начинается с описания подходов к классификации исследований и далее развивает идею эмпирического вопроса, намеченную в главе 1. Вы узнаете, как разрабатывать возможные исследования на основе ежедневных наблюдений поведения, теорий и вопросов, оставленных без ответов другими исследованиями. Глава завершается обсуждением *PsycINFO* — основной информационной базы данных по психологии и практическими советами по ее использованию. Изучив данную главу, вы сможете:

- понять разницу и определить значение: а) фундаментальных и прикладных исследований, б) лабораторных и полевых исследований и в) качественных и количественных исследований;
- понять, почему правильный эмпирический вопрос предполагает использование операциональных определений;
- привести примеры исследований, разработанных на основе ежедневных наблюдений и случайных открытий;
- дать определение и охарактеризовать особенности теоретического знания в психологии и показать: а) как теории приводят к эмпирическим исследованиям, б) как результаты исследований воздействуют на теории и в) что теории должны быть продуктивными, простыми и способными к фальсификации;
- понять важность вопроса «что дальше?» и значение исследований, одновременно повторяющих и углубляющих оригинальное исследование;
- объяснить роль креативного мышления в науке;
- использовать базу данных *PsycINFO* для поиска информации о психологических исследованиях.

На занятиях по исследовательским методам или в качестве самостоятельной работы вас могут попросить придумать идею для исследования. Вы можете ощутить, как свет меркнет у вас перед глазами, и испытать приступ паники. Держитесь, эта глава появилась как раз вовремя. Вряд ли идеи сразу начнут быстро и легко

приходить вам в голову, но по крайней мере вы поймете, с чего начать. Но прежде, чем описывать источники идей для исследований, давайте рассмотрим классификацию психологических исследований.

Виды психологических исследований

Есть различные классификации психологических исследований. Например, можно выделить фундаментальные и прикладные исследования, можно классифицировать их по условиям проведения и по преобладанию качественного или количественного анализа.

Фундаментальные и прикладные исследования

Исследования, посвященные описанию, предсказанию и объяснению фундаментальных законов поведения, называются **фундаментальными исследованиями.** Противопоставленные им **прикладные исследования** носят такое название, поскольку непосредственно связаны с решением конкретных проблем. Для иллюстрации различия между ними представьте себе исследование, посвященное изучению памяти. В фундаментальном исследовании будет изучаться структура памяти, участники будут запоминать список слов, воспроизводить его, изучать список еще раз, снова воспроизводить и так далее несколько раз (см., к примеру, Tulving, 1966). Идея такого исследования состоит в том, чтобы проверить, будут ли эти слова по ходу эксперимента запоминаться в одном и том же порядке, выявляя тем самым, как слова группируются в памяти участников. Такое исследование не имеет прямого практического применения, но может проводиться просто для изучения структуры памяти. Результаты такого исследования, вероятно, расширят знания о механизмах ее работы. Примером прикладного исследования памяти может быть эксперимент, при котором участники просматривают видеозапись несчастного случая, а затем стараются вспомнить все, что они увидели (см., напр.: Loftus & Palmer, 1974). Это исследование может иметь прямое отношение к важному для разработки законодательства вопросу о снятии показаний очевидцев.

Иногда считают, что прикладные исследования имеют большее значение, чем фундаментальные, так как они направлены на вопросы первостепенной важности. На это можно возразить, что основное преимущество фундаментальных исследований состоит в том, что общие законы можно применять в различных практических ситуациях. Тем не менее фундаментальные исследования очень часто представляют собой мишень для политиков, неистовствующих по поводу неправильного использования налогов на финансирование не слишком «полезных» исследований (деньги от налогов распределяются через гранты федеральных служб, в частности Национального научного фонда). Такое обвинение легко выдвинуть, и оно легко находит отклик у избирателей, ведь основная черта американского характера — высокая оценка прежде всего практически полезного. Например, президент Американского психологического общества — известный психолог-экспериментатор Ричард Ф. Томпсон в интервью, данном после своего избрания признал, что «многие из нас, тех, кто сегодня занимается фундаментальной наукой, чувствуют необходимость оправдать свое существование и стать полезными для решения проблем общества».

Конечно, фундаментальные исследования никогда не будут прекращены. Многие, если не все, прикладные разработки зависят от прочного основания, заложенного фундаментальными исследованиями. Без него идеи некоторых прикладных проектов никогда бы не возникли, а их реализация не была бы такой полной. Хороший пример тому — исследование чтения Иглендом в 1975 г. Целью эксперимента была оценка методики обучения дошкольников сходным буквам (например, *R* и *P*). В соответствии с методикой детям показывали карточки, аналогичные изображенным на рис. 3.1, и просили из шести букв выбрать такие же, как буква, изображенная вверху карточки. Игленд выделил отличительные особенности букв (например, «ножка» у буквы *R*, которая отличает ее от *P*), напечатав их красным. В течение нескольких попыток красный постепенно заменялся на черный. В сравнении с участниками, получившими только черные буквы, члены экспериментальной группы делали меньше ошибок. Они также лучше справились с дальнейшими тестами, проведенными через неделю.

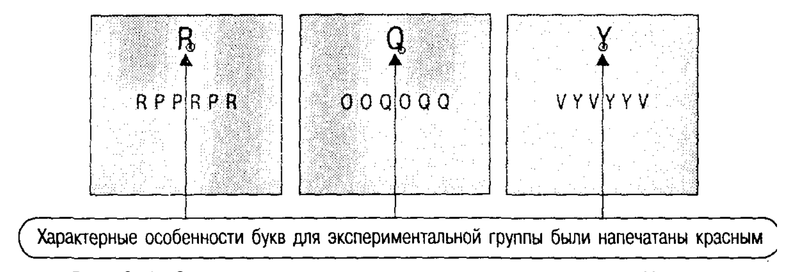


Рис. 3.1. Стимульные карточки, аналогичные использованным Иглендом

Нам важно отметить что в основе исследования Игленда лежало предположение, что на распознавание букв может влиять либо восприятие формы отдельных элементов, либо характерные особенности стимула. В то время когда Игленд проводил свои эксперименты, концепцию распознавания в основном формировала теория особенностей, поэтому многие фундаментальные исследования были посвящены изучению различных аспектов этой теории. Например, в одном из ранних исследований данного вопроса, проведенном Найссером (Neisser, 1963), участники просматривали массивы букв, аналогичные изображенным на рис. 3.2. Им необходимо было подать сигнал, как только они распознают нужную букву. Как видно из массивов, Найссер варьировал степень сходства характерных особенностей букв. Из-за большего сходства буквы *O* с буквами, похожими на *Q*, чем с буквами, похожими на *X*, участники дольше распознавали букву О, окруженную буквами *Q, U,S и G,* чем ее же, окруженную *X, A,N и* .

Хотя в своей работе с дошкольниками Игленд никогда не упоминал исследование Найссера или подобные ему, очевидно, что основа, созданная исследованиями, посвященными изучению теории особенностей, сыграла важную роль при разработке обучающей программы по чтению. Более того, в разработку этой программы значительный вклад внесло еще одно независимое направление фундаментальных исследований. Для изучения формирования у животных условных рефлексов была разработана процедура, названная «безошибочной» тренировкой дифференциации (например, Terrace, 1963), в которой использовалось такое же постепенное изме-

нение стимулов, как в тренировочной программе Игленда. Подобная ситуация возникает очень часто: в фундаментальных исследованиях психологические законы изучаются независимо, исключительно ради получения сведений о них, далее создается комплекс знаний о данных феноменах, а затем этот комплекс ложится в основу прикладных исследований, посвященных конкретным проблемам.

1. Найдите букву 0:

G	Q	Q	U
Q	S	G	G
U	Q	G	S
S	G	0	Q
U	Q	S	U
G	G	S	U

2. Найдите букву 0:

A	X	A	N
L	A	N	X
X	X	N	L
A	N	0	A
L	L	X	A
A	L	A	N

3. Найдите букву К:

G	Q	Q	U
Q	S	G	G
U	Q	G	S
S	G	K	Q
U	Q	S	U
G	G	S	U

4. Найдите букву К:

A	X	A	N
L	A	N	X
X	X	N	L
A	N	K	A
L	L	X	A
A	L	A	N

Рис. 3.2. Массив стимулов, использованный Найссером для изучения восприятия характерных особенностей в 1963 г.

Не только фундаментальные исследования часто приводят к прикладному изучению вопроса, но и результаты прикладных исследований, в свою очередь, нередко бывают важны для фундаментальных, подтверждая или опровергая выдвинутые теории. Подтверждение теории особенностей относительно распознавания паттернов не входило в задачи Игленда, но его исследование привело именно к этому. Аналогично описанное выше исследование памяти является прикладным, но оно внесло свой вклад в развитие общей теории долговременной памяти.

Условия проведения: лабораторные и полевые исследования

Другой подход к классификации исследований — разделение их на основании места проведения экспериментов. Как следует из названия данного раздела, исследования могут различаться в зависимости от того, проводятся ли они в лаборатории или вне ее. **Лабораторные исследования** обеспечивают исследователям высокую степень контроля: условия эксперимента могут быть определены более четко, а отбор и исследование испытуемых проводиться более систематично. С другой стороны, условия проведения **полевых исследований** ближе к ситуациям повседневной жизни. Хотя полевые исследования чаще бывают прикладными, а лабораторные — фундаментальными, необходимо помнить, что некоторые фундаментальные исследования проводятся вне лаборатории, а некоторые прикладные — в лабораториях.

Лабораторные исследования иногда критикуют за «искусственность» и оторванность от ситуаций повседневной жизни, но очевидно, что они стали источником важной информации о поведении, поэтому, чтобы судить о ценности исследования, нужно рассматривать более значимые его стороны, чем сходство его условий с повседневными. Например, социальный психолог Элиот Аронсон (Aronson, 1999a) различал жизненный и экспериментальный реализм. **Жизненный реализм** проявляется в точности отображения исследованием реальной жизни. **Экспериментальный реализм** проявляется в том, насколько лабораторное или полевое исследование «воздействует на субъектов, вынуждает их серьезно относиться к экспериментам и вовлекает в процесс исследования» (p. 411). По Аронсону, учитывать нужно именно экспериментальный реализм. Если участники вовлечены в исследование и воспринимают его серьезно, то исследователь сможет сделать правильные выводы о поведении. Степень жизненного реализма в проведенном Милграмом исследовании подчинения, которое обсуждалось в предыдущей главе, была невысока — маловероятно, что мы станем применять электрический шок к кому-то, кто не может запомнить список слов. Но очевидно, что отобранные Милграмом добровольцы были вовлечены в эксперимент, и поэтому исследование имело высокий уровень экспериментального реализма. Мы видели, что исследование Милграма оценивают неоднозначно, но ясно, что оно тем не менее пролило свет на факторы, влияющие на подчинение власти.

Сходство с реальной жизнью часто называют в ряду достоинств полевых исследований, но есть и другие причины для проведения исследований вне лаборатории. Например, Стернберг и Григоренко (Sternberg and Grigorenko, 1999) в 1999 г. на основании проведенных ими исследований когнитивного функционирования индийских детей, зараженных кишечными паразитами, указали несколько достоинств полевых исследований. Во-первых, условия полевых исследований часто не могут быть воспроизведены в лаборатории. Стернберг и Григоренко изучали детей в тесных кварталах, при 45-градусной жаре и в окружении почти невыносимого запаха экскрементов из открытой канализации. Такие условия довольно трудно создать в лаборатории, хотя бы потому, что исследовательский совет вряд ли разрешит подобное исследование. Вторая причина для проведения полевых исследований — необходимость подтвердить лабораторные исследования и по возможности скорректировать ошибки, вызванные естественными ограничениями лабораторных условий. Третья причина — возможность получения данных, способных быстро повлиять на жизнь исследуемых людей. В-четвертых, хотя полевые исследования обычно ассоциируются с прикладными, они также предоставляют широкие возможности для проведения фундаментальных исследований. Стернберг и его коллеги в разных местах земного шара изучали последствия заражения людей паразитами, и основным фокусом их работы была проверка гипотезы, построенной на основании теории Стернберга о природе интеллекта.

Некоторые исследователи объединяют лабораторные и полевые эксперименты в одно исследование. Хороший пример — проект Даттона и Эрона (Dutton & Агоп, 1974). Ученых заинтересовала проверка двухфакторной теории романтической любви, которая утверждает, что люди, испытывающие сильное физическое возбуждение, иногда принимают это возбуждение за любовь (два фактора, учитывающие-

ся данной теорией, — физиологическое возбуждение и его когнитивная интерпретация). Они поставили эксперимент, в котором мужчины сначала переживали страх различной степени, а затем встречались с привлекательной женщиной. Даттон и Эрон хотели проверить, будет ли вызванное страхом возбуждение расцениваться как отчасти связанное с физической привлекательностью женщины. Полевые эксперименты были проведены в двух разных участках канадского национального парка Британской Колумбии, где требовался переход через реку. Один из них — раскачивающийся подвесной мост длиной в 150 метров, располагающийся на высоте 75 метров над рекой (рис. 3.3), а второй — прочный деревянный мост, проходящий всего в 3,5 метра над водой. На обоих мостах привлекательная женщина подходила к мужчинам и просила помочь в проведении психологического исследования по выявлению воздействия живописных видов на креативность. Согласившимся участникам она выдавала мнимый тест на креативность и свой телефонный номер на случай, если у них возникнут вопросы по поводу проекта. По сравнению с испытуемыми на «спокойном» мосту, которые, скорее всего, почти не испытывали страха, мужчины на подвесном мосту показали в результатах теста большее количество сексуальных образов и чаще звонили участвовавшей в исследовании женщине.

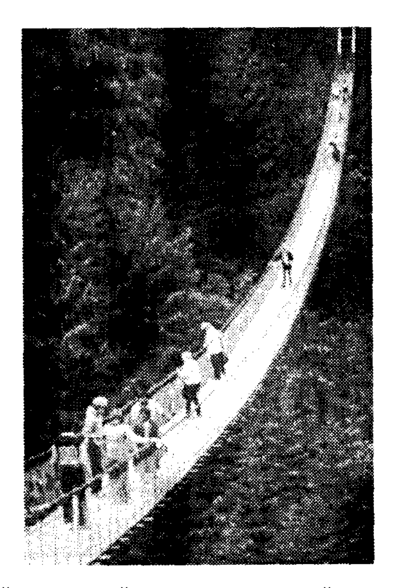


Рис. 3.3. Раскачивающийся подвесной мост, использованный в исследовании Даттона и Эрона

Результаты, полученные на подвесном мосту, соответствовали предположениям двухфакторной теории, но Даттон и Эрон совершенно справедливо решили, что их можно интерпретировать и иначе. Возможно, что мужчины, переходящие по подвесному мосту, были просто более смелыми и активными, чем другие, воспользовавшиеся деревянным мостом. Для проверки этого было проведено еще два эксперимента, один из них — в лаборатории. Даттон и Эрон набрали мужчин для изучения воздействия электрического шока на обучение и пригласили для помощи

в проведении эксперимента привлекательную женщину, которую все остальные приняли за обычную испытуемую. Одним участникам сообщили, что их подвергнут умеренному воздействию электрического шока, а другим — что они испытают сильный шок. Исследователи ожидали, что последнее сообщение вызовет более сильное физиологическое возбуждение. Именно так и произошло: Даттон и Эрон обнаружили, что мужчины, ожидающие воздействия сильного шока, испытывали более сильное физическое влечение к женщине, чем те, кто предполагал воздействие более слабого шока. Таким образом, лабораторный эксперимент подтвердил открытие, сделанное при полевом исследовании и состоявшее в том, что мужчины могут неверно интерпретировать возбуждение, вызванное страхом, как физическое влечение. Результаты обоих исследований подтвердили двухфакторную теорию любви.

Исследование Даттона и Эрона показывает, что полевые и лабораторные исследования могут привести к сходным результатам — это подтверждает, что оба вида исследований важны и необходимы. Но не являются ли результаты Даттона и Эрона случайностью? Можно ли сказать в целом, что результаты лабораторных экспериментов повторяют результаты полевых исследований? Конечно, да, по крайней мере в некоторых областях. Андерсон, Линдсей и Бушмэн (Anderson, Lindsay and Bushman, 1999) изучили некоторые вопросы из круга тем социальной психологии и обнаружили большое число лабораторных и полевых исследований (в целом 288) по изучению одинаковых переменных. Например, занимаясь областью изучения агрессии, они сравнили лабораторные и полевые эксперименты по изучению влияния анонимности на агрессивное поведение и обнаружили, что их результаты весьма похожи. Такой факт весьма важен для а) ученых, проводящих лабораторные исследования, уставших от обвинений в «искусственности» их экспериментов, и б) ученых, занимающихся полевыми исследованиями, уставших от обвинений в недостаточном контроле за экспериментами, который якобы не позволяет делать строгие выводы.

Еще один момент, который необходимо принять во внимание при выборе типа исследований, — это этика. Кроме того, что в лаборатории возможен более полный контроль за экспериментами, исследователи часто предпочитают лабораторные исследования из-за проблем получения осведомленного согласия и охраны частной жизни участников. В лабораториях относительно просто точно следовать стандартам этического кодекса, тогда как при проведении полевых исследований трудно, а зачастую невозможно, обеспечить осведомленное согласие участников и дебрифинг. Кроме того, в некоторых ситуациях процедура полевого исследования может рассматриваться как вмешательство в частную жизнь. В результате полевые исследования могут встретить большее недоверие со стороны исследовательского совета, а ученым потребуется представить более веские доказательства того, что важность их работы оправдывает риск участников. С другой стороны, как показывает пример исследования Стернберга и Григоренко, исследовательские советы могут не разрешить создание в лаборатории условий, аналогичных полевым. Прежде чем закончить изучение данной темы, прочитайте, пожалуйста, вставку 3.1 — в ней рассматривается вопрос вмешательства в частную жизнь с точки зрения закона.

ВСТАВКА 3.1

Этика — вопрос об охране частной жизни участников

В противоположность лабораторным условиям,во время полевых исследований иногда возникают проблемы с обоснованным согласием, возможностью выхода из эксперимента, дебрифингом и с ограничением вмешательства в частную жизнь участников. В интересном исследовании, проведенном в 1975 г., Силвермэн продемонстрировал то, почему исследователи иногда не решаются на проведение полевых экспериментов. Он показал описание 10 публикаций результатов полевых исследований двум адвокатам и спросил, не нарушены ли при их проведении какие-либо законы и не произошло ли вмешательства в частную жизнь участников. В ходе экспериментов участники исследования намеренно падали в вагоне метро, чтобы увидеть, придут ли окружающие на помощь, оставляли автомобили в различных местах, чтобы посмотреть, будут ли они разграблены, примеряли в обувном магазине множество пар обуви и просили мелочь у прохожих.

Два адвоката дали почти *противоположные* ответы. Первый адвокат считал, что оценивать исследования следует по их устремленности к получению полезного результата. Эксперименты были направлены на увеличение нашего знания о законах человеческого поведения, а не на достижение учеными личной выгоды. Он считал, что если бы против психологов были выдвинуты обвинения, то судья «стал бы сравнивать степень причиненного беспокойства с разумностью поставленных целей» (Silverman, 1975, p. 766). Второй адвокат, напротив, считал, что в некоторых экспериментах было достаточно оснований не только для гражданского процесса, поскольку некоторые люди, возможно, не хотели быть объектами исследования (вторжение в частную жизнь), но также для уголовного преследования на основании имевших место оскорблений, мошенничества, других криминальных действий и даже нарушения общественного порядка!

Силвермэн был настолько сбит с толку противоречивыми ответами, что представил описание проведенного в метро эксперимента, в котором исследовалась готовность людей помогать окружающим, судье с целью узнать его мнение о том, могут ли здесь быть выдвинуты гражданские или уголовные обвинения. В целом, судья согласился с первым адвокатом, по крайней мере в вопросе об уголовных обвинениях, но указал, что полевые эксперименты могут повлечь непредвиденные последствия, которые вызовут обвинения в халатности и дальнейшее судебное разбирательство. Таким образом, психологи, намеревающиеся провести полевое исследование, встречаются с трудностями, которых не возникает в лаборатории.

Вам, может быть, будет интересно узнать, что первый адвокат, не нашедший повода для судебного разбирательства, успешно занимался защитой по уголовным делам и его клиенты часто бывали оправданы. Второй адвокат специализировался в медицинском законодательстве и обычно «защищал законные права пациентов и объектов медицинских исследований» (Silverman, 1975, p. 767). В его понимании «вторжение исследователей в частную жизнь участников» было равно «причинению врачом вреда пациенту».

Количественные и качественные исследования

Большинство психологических исследований по природе своей количественные. В **количественных** исследованиях данные собираются и представляются в виде чисел — средних оценок для различных групп, выполнивших одно и то же задание, процентов людей, поступивших так или иначе, и т. д. В последние годы, однако, многие психслоги-исследователи стали проводить **качественные исследования,**

нередко заимствуя подходы у социологов и антропологов. Не так легко дать определение качественным исследованиям, но можно отметить, что они часто включают сбор подробной информации с помощью опроса отдельных людей или фокус-групп, в их составе иногда проводятся подробные исследования конкретных случаев, а также фундаментальные исследования методом наблюдения. Объединяет эти виды качественных исследований то, что их результаты представляются не как статистические отчеты, а в виде обобщающего анализа проекта. Многие исследования соединяют количественный и качественный подходы, и хороший пример этому — исследование, проведенное Уолкер в 1996 г. Его целью было выяснить, влияют ли различия в поведении мужчин и женщин при совместном просмотре телевизора на взаимоотношения супружеских пар в целом. Метод исследования — неформализованное индивидуальное интервью, в котором приняли участие 72 человека, составляющие 36 пар. Ответы на некоторые вопросы можно было обработать количественно. Например, вопрос, кто контролирует пульт, когда оба партнера смотрят передачу, показал, что женщины заняты этим только 20% времени. Большая же часть статьи была посвящена качественному анализу исследования, включала открытые вопросы, использованные в интервью, и цитаты из ответов участников, иллюстрирующие выводы. Например, Уолкер пришла к заключению, что хотя при совместном просмотре телевизора мужчины завладевают контролем над переключение каналов, источником конфликтов при этом обычно является решение вопроса о проведении свободного времени. Более подробно качественные исследования описываются в главах 10 и 12.

Постановка эмпирических вопросов

Вне зависимости от того а) посвящено исследование изучению фундаментальных или прикладных проблем или б) проводится в лаборатории или в полевых условиях, оно всегда начинается с постановки вопроса. Как вы помните из главы 1, я обозначил эти вопросы как эмпирические. Такие вопросы имеют две важные особенности: ответом на них должны быть факты, а все используемые термины должны быть точно определены.

Из главы 1 мы увидели, что по поводу таких очень интересных вопросов, как «Добры или злы люди по своей природе?» или «Существует ли личный Бог?», каждый может прийти к собственному выводу. Также возможно убедить остальных в своей правоте с помощью метода априори (по терминологии Пирса), но невозможно для ответа на них привести эмпирические данные. Следует отметить, однако, что возможны эмпирические вопросы о добре, зле и религии. Например:

- Какова взаимосвязь между верой в Бога и страхом смерти?
- Влияет ли вера в Бога на уровень болевого порога смертельно больных пациентов?
- Как влияет наличие альтруистически настроенного сиблинга на стремление стать донором крови?

Обратите внимание на то, что каждый из этих вопросов позволяет тем или иным способом собрать данные для получения ответа. Но прежде, чем производить сбор данных, необходимо произвести дальнейшее уточнение содержания вопросов. Эту задачу можно назвать задачей «операционализации» терминов вопроса. Процесс точного определения терминов — это вторая особенность постановки эмпирических вопросов.

Операциональные определения

Термин **операционизм** впервые появился в 20-х гг. XX в. в работе гарвардского физика Перси Бриджмэна «Логика современной физики» *(The Logic of Modem Physics,* 1927). Бриджмэн указал, что научная терминология должна быть абсолютно объективной и точной, а все понятия должны быть определены в терминах выполняемых операций. Такой вид определений был назван **операциональными определениями.** Длина некоего объекта, к примеру, может быть операционально определена с помощью набора установленных заранее процедур. По мнению Бриджмэна, «понятие длины тогда определено, когда установлены операции, с помощью которых измеряется длина, — это означает, что понятие длины включает в себя не больше и не меньше чем набор операций» (Bridgeman, 1927, p. 5).

Учитывая тенденцию психологов-экспериментаторов заимствовать подходы из более «взрослых» наук, неудивительно, что операционизм был принят психологическим сообществом, как только появился. Но строгий операционизм долго не просуществовал в психологии, отчасти потому, что уравнивание понятия и набора операций накладывает слишком сильные ограничения на понятие. Для психологов проблема операционизма свелась к вопросу о том, как реализовать его на практике, имея дело с такими сложными психологическими явлениями, как агрессия, креативность, депрессия и т. д. Физикам не так трудно прийти к согласию, принимая набор процедур для измерения длины линии, но как можно представить в виде операций, к примеру, понятие «агрессии»? Даже если социальные психологи согласятся с тем, что «агрессия — это поведение, выражающее намерение причинить вред» (Aronson, 1999a), какие именно действия должны измеряться в соответствии с данным пониманием? В литературе по агрессии этот термин был операционально определен как тип поведения, варьирующийся в пределах от воздействия электрошоком, сигналов автомобилиста и до затруднения выполнения необходимой задачи путем нажатия кнопки. Разве возможно здесь использование одного вида измерений?

Несмотря на проблематичность строгого применения операциональных определений, теория операционизма сыграла свою роль в психологии и научила исследователей четко определять используемые термины (Hilgard, 1987). Это особенно важно, если принять во внимание, что большинство психологических исследований использует понятия, для которых можно привести множество определений. Например, представьте себе исследование, посвященное изучению влияния голода на поведение в лабиринте. «Голод» — это термин, который может относиться к разным явлениям. Его нелегко определить для крысы: когда можно утверждать, что крыса голодна? Решить проблему можно с помощью операционального определения: выразить понятие в терминах процедуры (не кормить крысу в течение 24 ча-

сов — разумно предположить, что такие действия вызовут голод) или в терминах поведения (создать ситуацию, в которой крыса вынуждена усиленно работать, чтобы получить еду, — разумно предположить, что сытая крыса не станет выполнять задание).

Точность операциональных определений имеет одно важное следствие — они обеспечивают возможность воспроизведения экспериментов. Как вы узнали из глав 1 и 2, возможность повторения — это важная особенность любого научного исследования. Ограничения, налагаемые необходимостью точного определения терминов, не слишком удручают психологов-исследователей, так как требование точности увеличивает уверенность в надежности теорий поведения. Психологи также используют понятие **сходящихся операций,** означающее ряд исследований, использующих несколько отличные операциональные определения и экспериментальные процедуры, которые тем не менее приходят к одинаковым выводам; в итоге наше понимание механизмов поведения растет. Так, если отличные эксперименты по изучению влияния голода на обучение прохождению лабиринта приводят к одному и тому же выводу (даже если использовались разные операциональные определения голода и научения), то будет высока уверенность в том, что найден закон соотношения голода и успешности научения.

Умение ставить точные эмпирические вопросы требует некоторой тренировки техники постепенного сведения общих тем к конкретным вопросам. Эмпирические вопросы могут иметь разные основания: они могут происходить из а) ежедневных наблюдений над поведением, б) необходимости решить конкретную проблему, в) желания подтвердить или опровергнуть теорию или г) других оставшихся без ответа вопросов. При разработке идей исследования особенно полезно обладание навыками креативного мышления.

Разработка исследования на основе целенаправленного и случайного наблюдения поведения

Все мы имели возможность наблюдать чьим-либо поведением и задаваться вопросом о его причинах. Почему Норм так злится, когда промахивается мимо лунки, а Джефф, промахиваясь не реже, не обращает на это внимания и спокойно рассказывает о том, как ему повезло: не пришлось присутствовать на общем собрании. Почему дедушка помнит подробности Второй мировой войны, но не помнит, что он делал вчера? Почему одни студенты охотно помогают преподавателям проводить занятия, а другие не хотят и думать об этом? Почему одни маленькие дети очень общительны, а другие, из той же самой семьи, так болезненно застенчивы? И многое тому подобное.

Когда такие вопросы возникают у психологов-экспериментаторов, они становятся основой эмпирических вопросов. Для Роберта Стернберга, известного своими исследованиями типов интеллекта и человеческой любви, наблюдения за обычной жизнью становятся иточником вдохновения:

Все (или почти все) мои идеи возникли из наблюдения за людьми: за собой, за студентами, с которыми я работаю, за моими детьми, за моими отношениями с людьми, за отношениями других людей и т. д.... Дело в том, что в психологии нет лучшего источника данных, чем люди вокруг вас. Я всегда считал, что книги, лекции или лабораторная работа не больше способствуют формированию идей, чем живой опыт.

Р. Дж. Стернберг, частная беседа. 18 мая, 1993 г.

Одно из самых известных психологических исследований родилось именно так. Русский психолог Блюма Зейгарник, тогда студентка известного немецкого психолога Курта Левина (см. главу 8), раздала 164 участникам своего исследования от 18 до 22 простых заданий, выполнение каждого из которых требовало несколько минут. Заданием могло быть, например, изготовление картонной коробки, лепка глиняной фигурки, складывание картинок, а также решение арифметических и других умственных задач. Каждому участнику разрешали выполнить половину заданий, а затем прерывали и не позволяли закончить работу. Чтобы сделать остановку более заметной, Зейгарник «всегда выбирала момент, когда объект был наиболее погружен в работу» (1927/1967, р. 303). В результате исследования было обнаружено, что вероятность восстановления в памяти задания, на котором испытуемого прервали, была в два раза выше, чем для остальных заданий. Это явление (что незаконченные задания запоминаются лучше, чем законченные) сегодня называется «эффект Зейгарник».

Идея исследования возникла из обычного для аспирантов занятия — обсуждения исследований в кафе. Левин и его студенты часто встречались и проводили часы в кафе через дорогу от их лаборатории в Берлине. Они не могли не заметить, как один из официантов запоминает, что заказал каждый из студентов, не записывая этого. Однако вскоре после оплаты счета он полностью забывал о заказах. Может быть, до оплаты счета ситуация была «не завершена» и официанту требовалось удерживать информацию в памяти? Зейгарник очень заинтересовалась этим вопросом, был разработан эксперимент, а остальное, как говорят, дело техники.

Более новый пример того, как наблюдения приводят к исследованию, — эксперимент но социальной психологии, посвященный изучению готовности людей оказать помощь. Отправной точкой для исследования стали нашумевшие случаи отказа помочь. Самым известным из них был случай Китти Геновезе, имевший место в 1964 г. в Нью-Йорке: на женщину напали, а затем убили на глазах у 38 свидетелей, из которых ни один даже не попытался сделать анонимный звонок в полицию. Джон Дэрли, один из ведущих исследователей в области альтруистического поведения, позже вспоминал:

Без сомнения, катализатором для нашего исследования явилось описанное в «Нью-Йорк Тайме» убийство молодой женщины в Нью-Йорке, сегодня известное как случай Китти Геновезе. Молодая женщина была убита, но это, как ни прискорбно, обычный инцидент. Необычно в этой истории то, что тридцать восемь соседей смотрели в окна на происходящее и ни один из них не пришел на помощь. Бибб (Латане, коллега Дэрли) и я как-то ужинали вместе вскоре после этого события. Все вокруг говорили об этом и мы тоже... Думаю, что именно в тот день мы набросали на скатерти план экспериментов.

Krupat, 1975, р. 275

Случай Китти Геновезе привел Дэрли и Латанс к замыслу серии экспериментов, которые показали, что дело не в том, что очевидцев не волнует происходящее. Они часто думают, что если вокруг есть другие люди, то помочь может кто-нибудь из них (Darley & Latane, 1968). Сегодня эта тема хорошо исследована: обратитесь к любому современному учебнику по социальной психологии — он без сомнения посвятит данной теме целую главу.

К исследованию также может привести **случайное открытие** — обнаружение некоторого феномена при поиске чего-то совершенно другого. Такие открытия явились причиной многих важных событий в истории науки. Это происходит, если ученый бьется над разрешением трудной проблемы и некое непредвиденное событие случайно даст ключ к разгадке. Случайное открытие может быть сделано при неправильном ходе эксперимента, например при отказе аппаратуры. Получение Скиннером кривых торможения, вызванное неисправностью аппаратуры (см. главу 1), — хороший пример случайного открытия. Другой пример — случайное обнаружение в мозге детекторов признаков. Эксперимент, в ходе которого оно имело место, был частью фундаментального исследования по распознаванию особенностей, описанного выше в этой главе. (Вставка 3.2 посвящена рассказу о рождении исследования Дэвида Хьюбеля и Торстена Уиселя, получивших за него Нобелевскую премию.)

Исследование может быть не только результатом наблюдений обычной жизни, но и быть разработано специально для решения насущных проблем. Особенно это касается прикладных исследований, главной темы главы 10. Здесь мы ограничимся одним примером. Для облегчения адаптации студентов на первом семестре колледжи организуют специальный семинар «по выживанию». Администрация колледжа может провести прикладное исследование, сравнив группу первокурсников, посещающих экспериментальный семинар, с другой группой, проходящей обычную программу первого курса. Результаты исследования должны показать, стоит ли проводить этот семинар в будущем.

Разработка исследований на основе теорий

В главе 1 были кратко описаны цели науки психологии и в качестве одной из таких целей было названо объяснение поведения. По сути, процесс объяснения представляет собой процесс построения и проверки теорий. В данном разделе мы поговорим о том, что такое теория, рассмотрим взаимоотношения между созданием теории и сбором данных, а также способы распознания полезной теории.

Определение теории

Теория в психологии — это набор логически непротиворечивых утверждений о поведении, которые: а) обобщают существующие эмпирические знания о данном явлении, б) организуют эти знания в точные утверждения об отношениях между переменными (т. е. законы), в) дают предварительное объяснение явления и г) служат основой для предсказания поведения. Сделанные предсказания далее проверяются экспериментально. Теория подлежит пересмотру в соответствии с результатами эмпирических исследований.

ВСТАВКА 3.2

История — случайные открытия: эффект края

Одно из самых важных исследований второй половины XX в. по психологии зрительной системы было инициировано случайным открытием, сделанным в лаборатории Гарвардского университета Дэвидом Хьюбелем и Торсеном Уиселем (Hubel & Wisel, 1959). Они изучали функционирование отдельных нейронов на различных участках зрительного пути с целью определить, могут ли нейроны активироваться в соответствии с определенными стимулами. Экспериментальная установка состояла из экрана, на который проецировались различные образы. Голова наблюдавшей их кошки была зафиксирована в определенном положении, а в отдельные клетки ее зрительной системы были имплантированы электроды. (Уже в начале 50-х гг. процедура была настолько точно разработана, что можно было изучить активность отдельного нейрона.)

Хьюбель и Уисель расчитывали, что нейрон будет возбуждаться в ответ на проецирование на сетчатку глаза кошки черных или белых точек. Их первые попытки были неудачными. «Положение контакта микроэлектрода по отношению к коре было настолько устойчивым, что мы могли следить за одной клеткой в течение приблизительно девяти часов. Мы перепробовали все, что только приходило нам в голову, чтобы заставить электрод возбудиться» (Hubel, 1998, p. 69).

Хьюбель и Уисель упорно продолжали эксперименты и через некоторое время сконцентрировались на одном определенном участке сетчатки. Их заинтересовал тот факт, что прохождение точки через эту область иногда вызывало возбуждение нейрона, но при этом далеко не всегда: «После приблизительно пяти часов борьбы мы внезапно заметили, что стекло (слайд) с точкой иногда вызывает ответную реакцию, но было похоже на то, что эта реакция никак не связана с точкой. *Постепенно мы поняли: она вызывалась четкой, но слабой тенью, отбрасываемой краем стекла, когда мы опускали его в щель аппарата.* Вскоре мы убедились, что тень вызывает эффект, только если проходит через определенный небольшой участок сетчатки, и что край слайда должен занимать определенное положение в пространстве. Самым удивительным был контраст между пулеметной очередью, которой следовали реакции при правильном положении стекла, и полным их отсутствием, если мы изменяли положение или просто освещали яркой вспышкой глаза кошки» (Hubel, 1988, p. 69-70; курсив мой. - *Дж. Г.).*

Неожиданное открытие того, что существуют особые клетки зрительной системы, реагирующие на края и контуры, имеющие определенное положение в пространстве («детекторы края»), было только началом. Хьюбель и Уисель разработали расширенную исследовательскую программу определения разновидностей стимулов, приводящих в действие клетки на всех уровнях зрительного анализатора. Это открытие в 1981 г. принесло ученым Нобелевскую премию. В их работе отразилась та страсть ученых к проведению исследований, о которой мы говорили в главе 1 в связи с Гибсон и Скиннером. Рассказывая о годах исследований зрительных рецепторных полей (примерно с 1950 по 1980 г.), Хьюбель писал: «Я считаю, мне повезло, что я пережил эту эпоху радостных волнений. Некоторые эксперименты были очень тяжелыми, или они часто казались такими к 4 утра, особенно если шли неудачно. Но 98% времени мы были захвачены работой. Нейрофизиологические эксперименты дают моментальный эффект: можно сразу наблюдать ответную реакцию клетки на использованный стимул, а часто и одновременно понимать, какие функции мозга обеспечивают эти реакции» (Hubel, 1988, p. vii).

Теории различаются в зависимости от широты охвата материала. Некоторые из них стремятся наиболее широко охватить разные стороны поведения — примером является теория стадий Эриксона, объясняющая развитие и функционирование личности в различные периоды жизни. Но чаще теории более прицельно сфокусированы на отдельных аспектах поведения. Например, в социальной психологии теория справедливости направлена на объяснение взаимодействия людей в терминах вознаграждений, затрат и справедливости. В анормальной психологии с помощью теории выученной беспомощности пытаются объяснить психологическую депрессию. Теории также различаются по степени точности: некоторые из них имеют точное математическое выражение, другие описаны более просто и представляют собой набор логически связанных утверждений.

Для иллюстрации того, как возникают и развиваются теории, а также их важных особенностей, рассмотрим более подробно пример теории выученной беспомощности. Эта теория была выведена из исследования научения животных, а затем применена к проблеме депрессии у людей. Она является примером случайного открытия: приобретенная беспомощность была впервые обнаружена экспериментально совершенно неожиданно, в ходе исследования, преследовавшего абсолютно иные цели.

Это произошло во время работы аспирантов Расселла Лифа и Брюса Овермира в лаборатории Ричарда Соломона в Пенсильванском университете (Peterson, Maier & Seligman, 1993). Они проверяли предсказания, сделанные на основании теории обусловливания избегания, в подробности которой мы не будем вдаваться. В целом, они хотели перенести классическую условную реакцию на оперантную ситуацию избегания. Сначала они выработали у собаки условно-рефлекторный страх музыкального тона, сочетая его с электрическим разрядом. Использовалась предложенная И. П. Павловым процедура обусловливания, при которой собака удерживается в фиксированном положении похожим на сбрую устройством, названным «станок Павлова». Звучал сигнал, и подавался короткий электрический разряд. Через некоторое время сигнал превратился в условный раздражитель и собака начала его бояться. После первичного обусловливания Овермир и Лиф поместили собаку в «космический ящик» — камеру с небольшой перегородкой посередине. Они собирались научить собаку прыгать со стороны А на сторону Б при получении электрического разряда на стороне А, а также избегать разряда, перепрыгивая на противоположную сторону, при звуке сигнала, реакция на который уже была выработана. К их удивлению, когда собака получала разряд на стороне А, она не реагировала на это так, как обычно делают собаки (они начинают бегать и спасаются от разряда, случайно перепрыгнув через перегородку на сторону Б). Вместо этого, после непродолжительных хаотических движений, собака переставала сопротивляться, пассивно воспринимала разряд, часто ложилась на решетку и скулила (Overmier & Leaf, 1965).

Еще один аспирант Ричарда Соломона, Мартин Селигман, когда Овермир показал ему одну из «беспомощных» собак, очень заинтересовался этим явлением и начал изучать его более подробно. Была проведена серия экспериментов, направленных на выявление условий, при которых проявляется странное поведение. Например, на начальной стадии собака получала электрический разряд в станке

Павлова, хотела избежать его, но не могла, и постепенно прекращала попытки. Чем вызвана наступающая пассивность, самим повторяющимся разрядом или тем, что собака ничего не может с ним поделать (т. е. не может контролировать его)? Чтобы отделить эффект самого разряда от эффекта отсутствия контроля, Селигман и его коллеги провели несколько экспериментов, используя разработанный ими «триадный» план (Seligman, 1975). Одни собаки как и раньше получали электрический разряд, которым не могли управлять, другие получали такой же разряд, но могли отключить его, совершив некоторые действия, а третьи вообще не получали разряд[1]. Результат был недвусмысленным: беспомощность демонстрировали только собаки, перенесшие шок, которые не могли управлять получаемым разрядом (Seligman, 1975). Другие эксперименты показали, что собаки могут выработать «иммунитет» к беспомощности, если будут иметь возможность контролировать разряд, даже если затем потеряют эту возможность (Seligman & Maier, 1967).

Селигман назвал этот феномен «выученной беспомощностью». Он заключается в приобретенном нежелании избегать травматического воздействия после неоднократного повторения неудачных попыток контролировать его (Seligman, 1975). Собака понимала, что в станке Павлова шока невозможно избежать несмотря на все усилия, и прекращала попытки. Впоследствии, Селигман и его коллеги распространили свою теорию выученной беспомощности на человеческую психопатологию, а именно на теорию депрессии. С точки зрения теории выученной беспомощности депрессия развивается в результате переживания повторяющихся негативных событий, которых нельзя избежать или проконтролировать. Основные составляющие теории Селигман изобразил следующим образом (Seligman, 1975, p. 47):

Информация о возможной ситуации → Познавательная репрезентация возможной ситуации (научение, ожидание, восприятие, вера) → Поведение.

Процесс начинается с получения животным информации (через опыт) об отношениях между его поведением в определенных обстоятельствах и последствиями этого поведения (возможная ситуация). Селигман предположил, что выученная беспомощность может развиться, если животное поймет, что ситуация никак не связана с его ответной реакцией. Нормально животное переживает позитивные или негативные последствия своего поведения, а в данном случае его поведение совершенно непредсказуемым образом не приводит ни к каким последствиям. В результате формируется определенная «познавательная репрезентация возможной ситуации»: например, субъект начинает ожидать, что его действия не возымеют реальных последствий. Так, собака в станке Павлова, несмотря на разнообразные попытки вырваться из ремней, все равно получает электрический разряд. Ожидание этого приводит к закреплению поведенческой беспомощности. Селигман также предположил, что у людей выученная беспомощность может проявляться в виде психологической депрессии, так как после многих неудачных попыток контроля над последствиями своего поведения люди также могут прекратить их.

[1] Об этом виде исследовательского плана, плане с эквивалентной контрольной группой, вы узнаете в главе 7.

«Ожидание» в данной теории — это пример того, что психологи называют конструктом. **Конструкт** — это гипотетический фактор, который невозможно наблюдать непосредственно, его существование можно предположить на основании определенного поведения и определенных обстоятельств. Конструкт, названный «ожидание» — это основа теории Селигмана, в которой оно является «причиной мотивационного, познавательного и эмоционального истощения, сопровождающих беспомощность» (Seligman, 1975, р. 48). Наличие ожидания в данном смысле может а) быть логически выведено из факта прекращения животным попыток избежать отрицательного воздействия и б) предположительно иметь причиной повторение безуспешных попыток контроля за событиями.

Важным компонентом любой теории является ее способность к развитию на основе новых данных. Ни одна теория не абсолютна, и, как вы узнаете, прочитав еще несколько страниц, теория Селигмана — не исключение. Ход ее развития иллюстрирует взаимные отношения между теорией и фактами, а также демонстрирует важную характеристику хорошей теории — возможность предсказаний на ее основе, которые создадут почву для новых исследований. На этом необходимо остановиться подробнее.

Взаимосвязь теории и фактов

Переход от теории к фактам подразумевает применение логической **дедукции** — мышления от общих положений к частным событиям. В отношении к теории дедукция представляет собой вывод из допущения того, что (общая) теория правильна, того следствия, что (частное) событие *X* должно произойти с вероятностью выше вероятности случая. Предсказание конкретного события, сделанное на основании теории, называется **гипотезой.** Гипотезу можно определить как научное предположение того, что должно случиться при определенных обстоятельствах. Гипотеза ведет к разработке исследования, результаты которого соответствуют или противоречат сделанным предсказаниям. В первом случае теория подтверждается, во втором этого не происходит. Если теория подтверждена большим количеством исследований, уверенность в ее правильности высока. Другими словами, индуктивная база теории возрастает, если отдельные эксперименты дают все новые предсказанные результаты. **Индукция** — это логическое мышление от частного (конкретные результаты эксперимента) к общему (теория).

Конечно, эксперимент не всегда проходит так, как ожидается. Он может не осуществить правильной проверки гипотезы, например, если сделан неверный выбор операциональных определений изучаемых в эксперименте переменных, он может содержать методологические ошибки или совсем не подходить для данных целей. Кроме того, методы измерения психологических феноменов далеко не совершенны, поэтому неудачный эксперимент может быть результатом «измерительных ошибок» (более подробно вы познакомитесь с этим понятием в следующей главе). Поэтому один неожиданный результат редко ставит теорию под сомнение. Но если возникающие результаты раз за разом противоречат теории, особенно если они получены в разных лабораториях, доверие к этой теории ослабевает и она может быть отвергнута или, что более вероятно, существенно изменена.

Обратите внимание, что в двух предыдущих абзацах я избегал таких выражений как «сбывшиеся предсказания "доказывают" истинность теории» или «неудовлетворительные результаты "опровергают" теорию». Я поступал так потому, что ученые вообще стараются не использовать слова «доказывает» и «опровергает» при обсуждении теорий и фактов, имея на это как логические, так и практические основания.

С точки зрения строгой логики, невозможно доказать, что теория истинна и в то же время невозможно опровергнуть ее. Для понимания этого необходим краткий экскурс в правила формальной (если... то) логики. Предположим, что все известные вороны черные. Можно построить высказывание в форме условного предложения: «Если птица — ворона, то она точно будет черной». Далее предположим, что вы встречаете черную птицу. Можно ли заключить, что «следовательно, это должна быть ворона»? Нет, потому что другие виды птиц также могут быть черными. Заключить, что это — ворона, значит совершить логическую ошибку, известную как «утверждение результата». Ситуацию можно обобщить следующим образом:

Логическая ошибка «утверждение результата»:
Если птица — ворона, то она черная.
Эта птица черная.
Следовательно, это — ворона.

С другой стороны, предположим, вы видите желтую птицу. Можно ли заключить, что «следовательно, это не ворона»? Да, потому что дано, что все известные вороны черные. В логике этот тип заключения называется *«modus tollens»*. Таким образом:

Логически верный *modus tollens:*
Если птица — ворона, то она черпая.
Эта птица желтая.
Следовательно, это не ворона.

Различие между «утверждением результата» и *modus tolens* может непосредственно применяться к проверке теорий. Здесь высказывание «если ... то» примет форму: «Если теория *X* истинна, то можно ожидать событие *Y*». Обратимся еще раз к теории выученной беспомощности. Предположим, что я сделал следующее предсказание: «Если объяснение депрессии с помощью теории выученной беспомощности истинно, то процедура, снижающая депрессию у людей, должна также снижать уровень беспомощности у собак». Далее я разрабатываю эксперимент, в котором собакам с выученной беспомощностью будут давать антидепрессанты. Ход моей мысли таков: если эти лекарства помогают людям, находящимся в состоянии депрессии, и если причиной депрессии является выученная беспомощность, то лекарства, помогающие людям в состоянии депрессии, должны помогать собакам в состоянии беспомощности. Я провожу эксперимент и обнаруживаю, что собаки, принимавшие лекарства, стали вести себя нормально (т. е. перестали быть беспомощными). Такой результат действительно был получен в исследовании, проведенном Порсолом, Ле Пишоном и Жалфре в 1977 г. Если теперь я стану утверждать, что истинность теории выученной беспомощности доказана, то тем самым я совершу ошибку «утверждения результата»:

Если теория истинна, то антидепрессанты подействуют.
Антидепрессанты подействовали.
Следовательно, теория истинна.

На самом деле очевидно, что нельзя таким образом сделать вывод об истинности теории (т. е. о ее доказанности). Антидепрессанты могли подействовать по причине, совершенно не учитываемой теорией выученной беспомощности. Единственное, что можно сказать — и осторожные ученые никогда не скажут более того, — что эксперимент «подтверждает» теорию или «согласуется» с ней.

А что если антидепрессанты не подействуют и у животных не снизится уровень выученной беспомощности? С точки зрения правил логики, здесь применим *modus fallens,* и теорию можно считать ложной (т. е. опровергнутой);

Если теория истинна, то антидепрессанты подействуют.
Антидепрессанты не подействовали.
Следовательно, теория ложна.

Теперь, вспомните, пожалуйста, сделанное выше утверждение: что ученые обычно не говорят «доказывает» или «опровергает», и имеют на это не только логические, но и *практические* основания. Как мы видели, утверждать, что теория выученной беспомощности доказана, потому что лекарства подействовали, значит совершить ошибку «утверждения результата». Утверждение, что теория опровергнута, если лекарства не подействовали, формально правильно *(modus fattens),* но делать его неосмотрительно. Как отмечалось ранее, отдельные эксперименты по разным причинам могут дать результаты, не соответствующие предсказаниям, и ученые никогда не откажутся от теории из-за одного неудачного эксперимента. Даже серьезные опровержения могут не дискредитировать теорию полностью, а лишь указать границы ее применимости. Теории действительно отвергаются, но только если ученые теряют основания быть уверенными в них, только после многочисленных опровержений, полученных в нескольких лабораториях, и только когда конкурирующая теория начинает выглядеть более привлекательно.

Теории могут быть подтверждены или отвергнуты, но чаще всего по мере проведения исследований они развиваются. Леон Фестингер, создатель одной из самых устойчивых психологических теорий — теории когнитивного диссонанса (см. главу 12, вставку 12.1), — высказал следующую мысль об их судьбе: «относительно теорий важен не вопрос, могу ли я показать их истинность или ложность, но скорее важен вопрос, насколько полно они отражают эмпирическую действительность и как они должны изменяться, развиваясь» (Festinger, 1999, p. 383). Именно это произошло с теорией выученной беспомощности — она стала развиваться. Ряд экспериментов с участием людей привел к переформулированию теории и к тому, что она стала учитывать дополнительные когнитивные факторы (Abramson, Seligman & Teasdale, 1978). Например, в теорию вошел конструкт, названный «объяснительный стиль». Тип неудачника, склонный к пессимистическому объяснительному стилю, в большей степени демонстрирует выученную беспомощность, чем склонные к оптимистическому. Пессимисты «обвиняют себя и прогнозируют продолжение неудач на более длительный период и на большее количество

ситуаций» (Seligman & Schulman, 1986, p. 832). Оптимисты обвиняют не себя, а обстоятельства и не впадают в депрессию, даже если они разочарованы в такой же степени, как пессимисты.

Как это было и в исходном варианте теории выученной беспомощности, на основе новой версии делаются предсказания, которые должны проверяться на практике. Интересный пример такой проверки — полевой эксперимент со страховыми агентами (Seligman & Schulman, 1986). Представители именно этой специальности были выбраны в частности из-за того, что они сталкиваются с неудачами в среднем гораздо чаще. С помощью опроса определили типы способа объяснений и обнаружили, что у агентов с оптимистическим типом выше производительность труда (за два года они продали страховых полисов на 37% больше), а вероятность того, что они будут продолжать работать по этой специальности, в два раза превышает показатели их пессимистичных коллег.

Отличительные черты хороших теорий

История науки показывает, что одни теории оказывались более эффективными, чем другие. Более эффективные имеют некоторые отличительные особенности. Одна из самых их заметных особенностей — продуктивность. Хорошие теории умножают наши знания, продуцируя огромное количество исследований, — очевидно, эта черта присуща и теории выученной беспомощности. Две другие особенности хороших теорий — фальсифицируемость и простота — требуют некоторых пояснений.

Фальсифицируемость

Мнение о том, что психологические теории должны быть настолько хороши, чтобы они были способны объяснить все возможные результаты исследований, является распространенным заблуждением. В действительности теория, Которая претендует на объяснение всего, просто недоработана, и в очень серьезной степени. Чтобы понять, почему это так, необходимо рассмотреть подход к проверке теорий, впервые предложенный в 1959 г. теоретиком науки Карлом Поппером. Этот подход был учтен выше, когда мы обсуждали возможности доказательства и опровержения теорий.

Согласно Попперу, наука развивается в процессе создания теорий и дальнейших попыток их опровержения, или фальсификации. Устойчивые к постоянным попыткам фальсификации теории считаются возможно истинными (с ударением на «возможно»). Вспомните сделанное ранее замечание о том, что уверенность в правильности теории растет с ростом индуктивного подтверждения. Однако вследствие ограниченности метода индукции уверенность никогда не становится абсолютной. Например можно привести 100 примеров, которые обеспечат индуктивное подтверждение суждения «Все птицы могут летать», но достаточно одной нелетающей птицы (например, киви), чтобы разрушить его. Аналогично 100 сделанных на основании теории предсказаний могут поддержать ее, а одно опровержение может дискредитировать по принципу *modus tollens*. Конечно, как мы уже видели, на практике одно опровержение никогда не приводит к полному

отказу от теории, но тем не менее, с точки зрения Поппера, опровержение должно иметь приоритет над подтверждением или, по крайней мере, очень тщательно изучаться.

Если Поппер прав, то теории должны строиться таким образом, чтобы выведенные из них гипотезы в принципе могли быть опровергнуты. Должна существовать возможность получения экспериментального результата, способного опровергнуть гипотезу. Следствием этого являются требование известной точности в определении терминов, что было отмечено выше, а также принятие теорией возможности быть опровергнутой.

Как вы помните из главы 1, одна из характерных особенностей псевдонауки — это тенденция уклоняться от опровержений. Другими словами, псевдонаучные теории не проходят проверку на фальсифицируемость. Примером может служить френология — как вы помните, френологи приспосабливают свою теорию к объяснению (точнее, к оправданию) всех возможных отклонений, и таким образом создают впечатление ее безошибочности. Но, объясняя все, она по сути не дает возможности делать предсказания. Означает ли большая зона «жадности» у человека, что этот человек — вор? Согласно френологии, это возможно, но если жадность компенсируется большой зоной «скромности», то человек может и не быть вором. Такое рассуждение весьма неудовлетворительно.

Серьезный недостаток подхода Поппера заключается в том, что он не учитывает психологические особенности исследовательской работы — большинство исследователей в разгаре выполнения программы точно так же, как френологи, вырабатывают личное отношение к своим теориям и начинают искать подтверждающие свидетельства. Но в противоположность френологам и другим псевдоученым, настоящие ученые хорошо понимают важность стратегии фальсификации. Хотя исследователи и надеются найти подтверждение своих теорий, они всегда стараются ставить эксперименты, дающие возможность исключить то или иное объяснение опытного факта.

Обычная стратегия состоит в том, что в случае возможности нескольких объяснений ставится ряд экспериментов, в результате которых последовательно опровергается один из вариантов и подтверждаются остальные. Например, в исследовании Селигмана по изучению выученной беспомощности был разработан эксперимент с использованием триадного плана, который давал возможность опровергнуть предположение, что состояние беспомощности — это результат травмы, вызванной электрическим шоком. В двух группах собак, которые получали одинаковое количество электрических разрядов, не было выявлено признаков выученной беспомощности там, где собаки могли контролировать шок, тогда как в другой группе возник ожидаемый эффект. Такие результаты позволили Селигману отбросить (т. е. фальсифицировать) версию, предполагавшую, что состояние беспомощности зависит исключительно от количества полученного электрошока.

Исследование мнимых способностей одной лошади к математике и чтению — известный исторический пример применения стратегии фальсификации. Описание истории Умного Ганса — лошади с мнимыми интеллектуальными способностями — приведено во вставке 3.3.

ВСТАВКА 3.3

Классические исследования — фальсификация и Умный Ганс

В Берлине рубежа XX в. лучшее после только что открытого метро шоу проходило в одном из дворов, примыкающих к улице Гринбенау. Здесь зрителю демонстрировали лошадь (рис. 3.4), имеющую незаурядные интеллектуальные способности. Когда хозяин коня, Вильгельм фон Остен, просил его умножить 4 на 4, он 16 раз бил передним копытом. Выполнение задач на сложение, вычитание, умножение и деление не вызывало сложностей у замечательного коня. Даже простые и десятичные дроби не были проблемой. Когда его просили сложить $^{2}/_{5}$ и У , он отбивал 9 раз числитель и 10 раз знаменатель (Sanford, 1914). Конь также умел читать и делить слова на буквы с помощью системы, переводившей буквы в число ударов.



Рис. 3.4. Умный Ганс за работой

Если у вас уже сформировались навыки научного мышления, я думаю, вы скептически отнесетесь к существованию лошади, которая умеет читать, а вычисляет лучше, чем некоторые из ваших друзей. В то время также были скептики, и один из них, Оскар Пфунгст, дал замечательный пример использования стратегии фальсификации по Попперу. Пфунгст решил проверить, нельзя ли найти более разумное (т. е. более простое) объяснение поведения коня.

Специальная комиссия ученых и дрессировщиков заключила, что фон Остен не мошенничает. Но Пфунгст подозревал, что владелец может подавать коню незаметные сигналы, влияющие на его ответ. Он решил, что в таком случае конь будет отвечать правильно, только если задающий вопрос человек сам знает ответ на него. И тогда математические способности коня могут быть объяснены тем, что фон Остен - математик.

Проверить, будет ли конь знать ответ, если его не знает задающий вопрос человек, не составило труда. Пфунгст провел несколько экспериментов, в которых спрашивающий знал правильные ответы только на часть вопросов. Например, человек показывал коню карточку с числом, и, если ему разрешалось предварительно посмотреть на карточку, конь отбивал число

копытом правильно в 98% случаев. Если человек до ответа коня не видел карточку, то правильные ответы составляли только 8% (Fernald, 1984). Так были проверены математические способности. С помощью ряда похожих экспериментов Пфунгст опроверг (фальсифицировал) способность Ганса пользоваться языком.

Таким образом, стало ясно, что Ганс получает информацию о правильных ответах от задающего вопрос человека. Но оставалось загадкой, как он это делает. Со временем Пфунгст нашел ответ и на этот вопрос. Он установил, что конь реагировал на очень слабые визуальные сигналы, подаваемые задающим вопрос человеком. Когда кто-нибудь задавал вопрос, он слегка наклонялся вперед или опускал глаза, не осознавая этого (вероятно, глядя на копыто коня, ожидая, что он начнет стучать). Ганс расценивал такие движения как сигнал к началу ответа. Когда он отбивал правильное количество ударов, человек выпрямлялся или поднимал глаза - лишь слегка и не осознавая этого, но для Ганса этого было достаточно, чтобы остановиться.

Кроме стратегии фальсификации история Умного Ганса иллюстрирует еще две связанных с психологическим исследованием темы. Во-первых, показав, что способности коня свидетельствуют не о высоком уровне его интеллекта, а объясняются в пределах процесса научения ответной реакции на два вида визуальных сигналов (когда начинать и когда заканчивать), Пфунгст дал более *простое* объяснение поведения коня. Во-вторых, если фон Остен подавал незаметные сигналы, влияющие на поведение, то, вероятно, и экспериментаторы вообще могут незаметно воздействовать на поведение участников, предполагая, какие результаты должны быть получены. Это пример явления, известного как *искажение, вызванное экспериментатором.* Мы вернемся к нему в главе 6.

Простота

Хорошие теории, помимо того, что они потенциально фальсифицируемы, также обычно простые теории. В идеале это означает, что они включают минимальное число конструктов и допущений, необходимых для адекватного объяснения явления и предсказания результатов будущих исследований. Если две теории эквивалентны по всем показателям, кроме простоты, то предпочитается обычно более простая.

В психологии эту идею выдвинул Конви Ллойд Морган — британский психолог, живший в конце XIX в. Это было время, когда теория эволюции толкала натуралистов к поискам у животных признаков психических процессов (например, мышления у лошадей — история Умного Ганса) для подтверждения идеи Дарвина о непрерывности эволюционной цепи. Эти поиски породили множество преувеличений, например, заявления о том, что мотыльки летят на свет свечи из любопытства, что бобры, делая запруды, показывают дальновидность и способность к планированию, а муравьи «содержат домашних животных» (Romanes, 1886, р. 83). Морган высказал идею о том, что поведение животных должно объясняться как можно проще. Известный «закон Ллойда Моргана» гласит: «Мы не можем интерпретировать действие как результат применения высшей психической способности, если его можно считать результатом применения способности, занимающей более низкое положение на психологической шкале» (Morgan, 1903, р. 53). Например, вместо того чтобы приписывать собаке, поднимающей щеколду замка, способность к рассуждениям, Морган объяснял ее поведение более просто, как результат обучения методом проб и ошибок. Собака сделала много попыток выйти со двора и в конце

концов обнаружила этот способ. Этот тип поведения постепенно закрепился повторявшимся успехом, и животное научилось выходить со двора.

Хорошим примером служит сравнение фрейдистской и бихевиористской теорий, объясняющих подражание четырехлетних мальчиков отцам. Фрейдистское объяснение требует большого количества допущений и конструктов, в том числе идей детской сексуальности и влияния бессознательного на поведение, эдипова комплекса, боязни кастрации, вытеснения и идентификации с агрессором. Эта теория говорит, что маленький мальчик испытывает сексуальное влечение к своей матери, но боится быть кастрированным, если желание проявится. Поэтому он вытесняет желание в бессознательное и идентифицирует себя с агрессивным отцом. Теория научения в этом случае просто допускает, что а) подкрепляемый образец поведения имеет тенденцию появляться в будущем в сходных ситуациях и б) родители склонны замечать и поощрять подражание. Очевидно, что теория научения в данном случае представляется более простой, чем ее соперница и при этом дает адекватное объяснение явления, а также основу для предсказаний результатов будущих исследований.

Разработка исследований на основе других исследований

Большая часть этого раздела главы о разработке идей для исследований посвящена более подробному обсуждению рассмотренной ранее темы — отношений между теорией и фактами, хотя исследования, имеющие источником другие исследования, могут проводиться даже без предварительной разработки теории. Иногда исследователи просто хотят изучить некоторое явление с целью открыть постоянные, предсказуемые взаимосвязи переменных (например, открыть закон поведения) и не стремятся к созданию теории. Исследование Скиннера по оперантному обусловливанию (глава 11) относится именно к этой категории.

Я считаю, что наиболее распространенным источником идей для исследований по психологии являются вопросы, оставленные без ответов другими, недавно проведенными исследованиями. Психологи проводят не изолированные эксперименты, а строят **программу исследований,** включающую серию взаимосвязанных экспериментов. Вы не встретите ученого, ставящего эксперимент по изучению готовности людей оказывать помощь, а затем переключающегося на исследование агрессии. Напротив, исследователи занимаются определенной областью знаний и в ее пределах разворачивают серию экспериментов, которые могут продолжаться годы и привлечь других исследователей, интересующихся данной проблемой. Выводы одного проекта неизменно имеют следствием другой, так как отвечая на одни эмпирические вопросы, эксперименты ставят новые. Исследование Селигмана, его коллег и студентов, посвященное выученной беспомощности, — это хороший пример такой исследовательской программы.

Несомненное свидетельство того, что одни исследования влекут за собой другие, можно найти, просмотрев любой из последних номеров журнала по психологии. Взгляните на имена авторов конкретных публикаций, а затем проверьте, встречаются ли те же имена в прилагаемых списках литературы в ссылках на похожие

исследования. К примеру, первые три номера «Журнала по экспериментальной психологии: обучение, память и познавательные процессы» за 1992 г. содержат 52 исследовательские статьи. Авторы статей ссылаются на другие свои работы в 48 из 52 статьей. Хотя некоторые из ссылок можно считать проявлением естественного желания ссылаться на свои работы, большей частью это обилие автоссылок отражает тот факт, что исследователи просто не ставят единичных опытов, а проводят систематическую программу взаимосвязанных экспериментов. Эксперименты ведут к новым экспериментам.

Исследовательские группы и вопрос «что дальше?»

Если вы попросите психолога-исследователя описать свою повседневную жизнь, вы получите самые разнообразные ответы, но выясните одно общее правило: исследователи почти никогда не работают в одиночку. Напротив, в своих лабораториях они собираются в исследовательские группы. Обычно группы состоят из главного исследователя — доктора *X,* нескольких аспирантов, работающих с доктором *X,* и, возможно, одного или двух увлеченных студентов, убедивших доктора *X* в своем интересе и желании работать в свободное время, хотя бы чистить клетки животных. Такая группа будет ставить одновременно несколько экспериментов, а ее члены будут проводить долгие часы в лаборатории, собирая данные и анализируя их за чашкой кофе. Они часто будут собираться в ближайшей забегаловке, подобно Левину и его студентам, и обсуждать исследовательский проект на различных этапах его проведения (и снова пить кофе). При обсуждении завершенного проекта, они будут применять подход, который можно назвать «что дальше?»: что мы должны делать дальше, имея теперь результаты этого исследования? В какой-то момент обсуждения кому-то придет в голову идея, и он задаст самый частый в беседах психологов-исследователей вопрос: «Как вы думаете, что произойдет, если мы сделаем X?», «X» — это черновая идея исследования, а «как вы думаете, что произойдет?» — вопрос о предсказании его результатов. Вопрос, приведет к оживленной дискуссии, в которой группа уточнит идею или возможно отвергнет ее и станет думать над следующим вопросом «что делать?». Если идея принята, будет разработана процедура исследования, которая в ближайшие дни будет опробована в **пилотажном исследовании,** усовершенствована (на этом этапе снова пьют кофе) и в конце концов преобразована в строгий эксперимент, который и будет впоследствии проведен.

Закончившись, исследование редко остается без продолжения, и его результаты почти всегда приводят к другим экспериментам, нередко для выяснения оставшихся без ответов вопросов начального исследования. Для примера рассмотрим серию из двух экспериментов, посвященную распознаванию лиц (Burton, Wilson, Cowan, and Bruce, 1999) В первом эксперименте авторы выявляли способность участников узнавать людей, изображения которых были записаны с помощью обычной аппаратуры для наблюдений со стандартнонизким качеством. Исследователи сравнивали результаты участников, знакомых с людьми на изображении, с результатами других участников, увидевших их впервые. Третья группа состояла из офицеров полиции, проходивших обучение (тренировавшихся опознавать воров по записям, сделанным скрытыми видеокамерами во время ограбления банка).

Исследователи обнаружили, что участники, относительно знакомые с людьми на видеозаписи, довольно хорошо справились с заданием на узнавание, тогда как показатели участников двух других групп были весьма невысокими. Получив такие результаты и задумавшись над вопросом «что дальше?», Бертон и его исследовательская группа решили выяснить причину правильного узнавания людей на видеозаписи, в случае если участники знакомы с ними. В чем именно было дело, в лицах, фигурах или, может быть, походках людей с видеозаписи? Эти вопросы привели к соответствующему эксперименту, в котором записи были обработаны так, чтобы скрыть лица, фигуры или двигающиеся ноги людей на видео. Было установлено, что узнавание происходит довольно точно, если скрыты фигуры или походка, но если участники не видят лиц людей на видеозаписи, то точность падает. Второй эксперимент вытекал из первого и отвечал на поставленные им вопросы.

Таким образом, исследование в психологии: а) обычно состоит из последовательной серии взаимосвязанных экспериментов, б) часто является результатом совместных усилий людей, посвящающих свой труд одной и той же узко специализированной области и в) почти не имеет структуры на начальных стадиях работы. В 1958 г. отсутствие структуры было отмечено группой выдающихся психологов-экспериментаторов, собранных Советом по образованию и обучению *APA* для выработки рекомендаций к аспирантским занятиям по экспериментальной психологии. Они охарактеризовали «исследовательский процесс, а соответственно и процесс разработки самой науки психологии, (как) весьма неформальное, часто нелогичное и иногда беспорядочное дело. Он в значительной части состоит из барахтанья в эмпирическом мире, которое иногда определяют как "пилотажное исследование", или "пробное исследование"» (Taylor, Garner, & Hunt, 1959, p. 169).

Одна из недавних разработок подхода «что дальше?» — расширение исследовательской группы за пределы отдельной лаборатории. В наш компьютерный век ученые различных университетов часто общаются посредством электронной почты. Виртуальные беседы часто включают вопрос «как вы думаете, что произойдет, если мы сделаем вот что...» и последующие описания предлагаемых методов. Ученые, которых разделяют тысячи миль, тем не менее могут вести неформальные обсуждения, становящиеся затем исследованиями. Они даже могут пить кофе, не прерывая виртуальную беседу.

Повторение и дополнение

Многие эксперименты, следующие за другими опытами, близки последним достаточно, чтобы считаться повторениями, но в то же время отличаются достаточно, чтобы не быть их точными копиями. Другими словами, они объединяют повторение и дополнение. Психологи-исследователи обычно используют термин **повторение** в отношении экспериментов, полностью или частично дублирующих процедуры исходного эксперимента. **Дополнение,** в свою очередь, сходно с первичным экспериментом и частично повторяет его, но идет дальше и имеет по крайней мере одну особенность. Более того, при экспериментах-дополнениях нередко используется термин **частичное повторение,** означающий часть эксперимента, повторяющую какой-либо отрезок проведенной ранее работы. Иногда для описания пошагового воспроизведения эксперимента используются термины «точное повторение» или «прямое повторение».

В целях обучения процедура точного повторения использовалась в известной лаборатории И. П. Павлова в России. Когда в лабораторию приходил новый сотрудник, его первой работой было повторение некоторых поставленных ранее экспериментов (Babkin, 1949). Таким образом, Павлов создал систему непрерывной проверки результатов в ходе выработки у новых исследователей навыков для проведения исследований, дополняющих сделанные ранее открытия. Однако в общем точное повторение проводится редко, по той простой причине, что ученые теряют в карьерном росте, если вся их работа состоит в повторении того, что сделал кто-то другой. Обычно точное повторение проводится только в том случае, если по поводу некоторых открытий возникают серьезные вопросы. Например, если несколько исследователей пытаются развить определенное открытие и их эксперименты включают частичное повторение, которого не удается добиться, может возникнуть необходимость в возвращении к оригинальному эксперименту и его точном повторении, чтобы выяснить степень надежности сделанного открытия. Как вы помните из предыдущей главы, неудачные попытки повторения экспериментов также приводят иногда к обнаружению научного мошенничества.

Эксперимент Мэриена, Вернера и Кула (Marean, Werner and Kuhl, 1992) представляет собой хороший пример одновременных повторения и дополнения исследования. Ученых заинтересовал вопрос, различают ли двухмесячные дети гласные звуки. Эксперимент был дополнением проведенного ранее исследования, показавшего, что шестимесячные дети обладают такой способностью. Мэрией и др. хотели узнать, развита ли эта способность у детей младше шести месяцев. В ходе эксперимента они исследовали двух- и трехмесячных детей и как частичное повторение предыдущего эксперимента включили в число испытуемых также шестимесячных детей. В целом эксперимент показал, что уже в два месяца дети по-разному реагируют на два различных гласных, произнесенных одним человеком, но показывают одинаковую реакцию на один гласный, произнесенный разными людьми. Это означает, что они проводили различие на основе общих характеристик гласных звуков, а не по индивидуальным акустическим особенностям двух **разных** голосов.

Креативное мышление в науке

В этой главе несколько раз подразумевалась, но явно не обсуждалась одна составляющая процесса разработки исследований — научная креативность. Нетрудно говорить о том, что исследование может быть разработано на основе простых наблюдений, теории или результатов других исследований, но переход от источника идеи исследования к самому исследованию не произойдет автоматически. В какой-то момент эксперименты должны быть разработаны. Иногда эксперименты логически вытекают из предыдущей работы и почти не нуждаются в креативности, но бывает и так, что при их разработке требуется творческий подход.

Креативное мышление при разработке исследований включает процесс определения значимых связей между на первый взгляд бессвязными идеями и использование этих связей для создания экспериментов. Такое мышление не протекает

в вакууме, напротив, оно включено в контекст проблемы, которую решают ученые, обладающие немалыми знаниями о ней. Как выразился известный биолог Луи Пастер, «случай благосклонен к подготовленному уму» (процитировано в Myers, 1992, p. 335). Таким образом, случайность сама по себе не дает идею эксперимента — погруженный в проблему ученый должен увидеть случайное событие как недостающую часть, необходимую для решения поставленного вопроса. Это одна из причин, по которой исследователи работают в группах, ведь присутствие нескольких мыслящих людей увеличивает вероятность того, что кому-нибудь придет в голову идея, которая предстанет другому члену группы как недостающая деталь головоломки.

В качестве примера использования научной креативности рассмотрим исследования с использованием лабиринта. Попросите психолога назвать основные составляющие исследовательского оборудования, и первым (или близко к началу списка) он назовет лабиринт. Хотя эксперименты с лабиринтами достигли вершины своей популярности в 1920-1940-х гг., лабиринт все еще остается важным инструментом для исследования научения и пространственного ориентирования. Честь проведения первого эксперимента по изучению поведения крыс в лабиринте принадлежит Вилларду Смоллу из Университета Кларк, проводившему свои исследования в конце XIX в. (Small, 1900).

Как Смоллу пришла идея поместить крыс в лабиринт? Вместе со своим коллегой по лаборатории, Линусом Клайном, он изучал в основном поведение крыс, в частности их «способность к отысканию дома». В разговоре с Эдмундом Сэнфордом, директором лаборатории Кларка, Клайн описал увиденные им туннели, «прорытые большими дикими крысами к своим норам под крыльцом старого дома... Туннели находились на глубине от семи до пятнадцати сантиметров под землей и, открытые во время раскопок, представляли собой настоящий лабиринт» (Miles, 1930, p. 331). Вероятно, слово «лабиринт» замкнуло цепочку для Сэнфорда, и он предложил Клайну самому построить лабиринт, использовав в качестве модели Хэмптон Корт — самый популярный в Англии лабиринт размером в человеческий рост. К моменту разговора Сэнфорд только что вернулся из творческого отпуска, во время которого он ездил в Англию, где мог посетить Хэмптон Корт.

Имея другие незавершенные проекты, Клайн передал идею Смоллу, который построил из проволочной сетки лабиринт размером 180 на 240 см, изменив треугольную форму Хэмптон Корта (рис. 3.5, а) на квадратную (рис. 3.5, б) и оставив без изменения общую схему. Смолл провел несколько экспериментов, исследуя изучение крысами лабиринта. В первые десятилетия XX в. схема лабиринта Хэмптон Корт стала весьма популярной, а ее использование положило начало традиции изучения поведения крыс в лабиринте, существующей и сегодня[1].

[1] Хотя критики иногда указывают на лабиринты как на пример «искусственности» лабораторных исследований по психологии (т. е. отсутствия привычной для крыс реальности), стоит отметить, что исходной целью Смолла было создать не стерильные условия, а близкие к крысиному миру, или, словами Смолла, создать «как можно меньше различий между условиями эксперимента и обычными условиями жизни» (Small, 1900, p. 209).

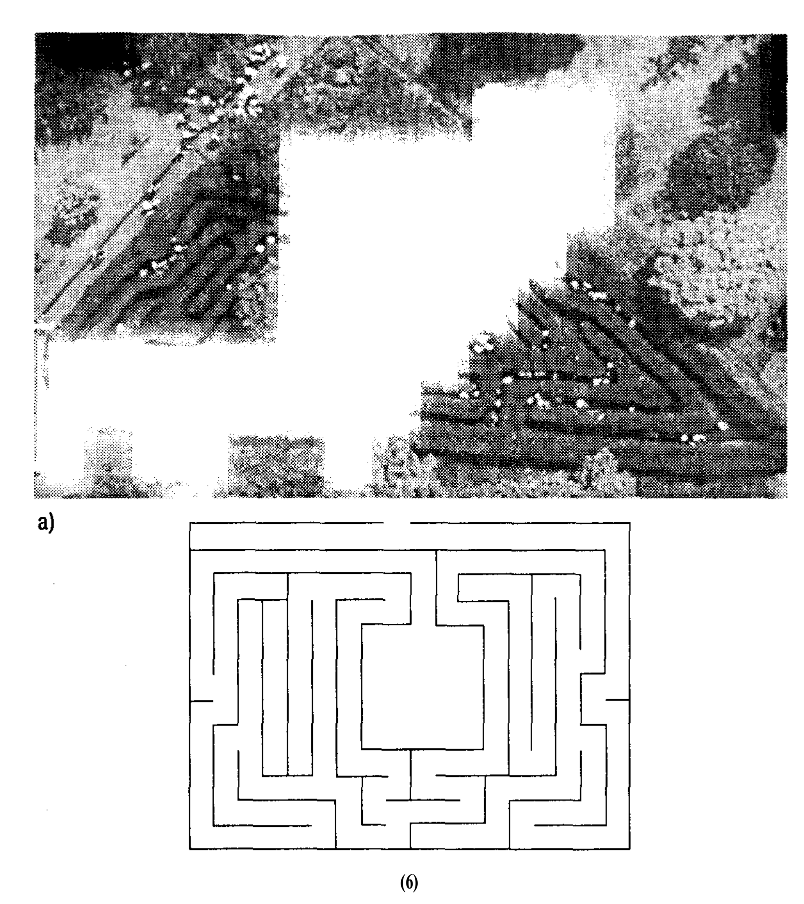


Рис. 3.5. а) Лабиринт Хэмптон Корт в оживленный день, б) Адаптация Смоллом схемы лабиринта Хэмптон Корт для эксперимента по изучению поведения крыс в лабиринте

Эта история хорошо иллюстрирует, как работает научная креативность. Ученые (Клайн и Смолл), обладающие знаниями в определенной области исследований (поведение животных) боролись с трудной проблемой (каким образом исследовать способность крыс к отысканию дома). Случайное замечание Клайна (рассказ о крысиных туннелях под крыльцом), соединившись со знанием Сэнфорда о лабиринте Хэмптон Корт, обнаружило связь между двумя на первый взгляд несвязанными событиями, и проблема была решена — для изучения способности крыс к отысканию дома необходимо создать устройство, копирующее известный английский лабиринт.

Следует отметить, что хотя основательные знания в определенной области являются предпосылкой креативного мышления в науке, наличие этих знаний иногда повышает ригидность мышления и снижает креативность. Случается, что ученые настолько привыкают к определенному методу или теории, что не способны рассматривать альтернативные варианты, а следовательно, возможность новых открытий для них снижается. Вернемся к исследованиям с лабиринтами.

Лабиринты внесли значительный вклад в наше понимание фундаментального процесса научения и открытие их использования великолепно иллюстрирует феномен научной креативности. Однако лабиринты могут также «завести в тупик». Общепринятое устройство, будучи стандартно применяемо, может снизить креативность ученых и привести к сужению поля работы до ситуаций, возможных только в лабиринте. Исследование феномена «центробежного отклонения» — пример такого развития событий. В 1920-1930-х гг. данное явление тщательно исследовалось. Было выяснено, что животные при повороте в лабиринте (возможно, только при большой скорости движения), как правило, отклоняются центробежной силой к противоположной направлению поворота стене, что определяет направление следующего поворота. Такому отклонению противопоставляли «тенденцию к движению вперед». Чтобы выявить факторы, обеспечивающие либо отклонение, либо движение вперед, были проведены десятки экспериментов (например, Schneirla, 1920). Эксперименты были разработаны весьма изящно и без сомнения помогли многим психологам-экспериментаторам развить свои исследовательские навыки, но значение этого исследования ограничивалось функционированием самого лабиринта и оно не смогло пролить свет на фундаментальный процесс научения.

Вероятно, известный бихевиорист Э. К. Толмен был только наполовину серьезен, когда в 1937 г. закончил свою президентскую речь (он являлся президентом ассоциации *APA)* словами: «Все важные вещи в психологии... могут быть, в сущности, открыты с помощью... анализа детерминант поведения крыс на развилках лабиринта» (процитировано в Hilgard, 1978, p. 364). Это замечание показывает, как прибор может ограничивать научное мышление. История создания экспериментального оборудования может прекрасно проиллюстрировать работу научного мышления (например, идея Сэнфорда об использовании схемы лабиринта Хэмптон Корт), но новаторский потенциал заметно ослабляется, когда новые прибор или процедура исследования упрочат свои позиции. Обширные знания о лабиринтах и их экспериментальном применении могут помочь исследователям творчески подойти к разработке экспериментов, но они же могут и снижать креативность.

Обзор литературы

Исследовательские проекты не разрабатываются в вакууме. Психологи, вовлеченные в программу исследований, прекрасно осведомлены не только о работе своей собственной лаборатории, но также и об аналогичной работе, проводимой в других лабораториях. Ученые, разрабатывающие эксперименты на основе какой-либо теории, также знакомы и с другими исследованиями, базирующимися на данной теории. Даже исследователи, к которым идея эксперимента приходит в результате наблюдений, направленных на выявление причинно-следственных связей в некотором явлении, часто проводят такие наблюдения не изолированно, а в контексте определенной проблемы или соответствующей информации. Как же получить информацию о литературе по проблеме? Хорошим началом для знакомства с литературой является знакомство с компьютерными базами данных.

Поиск в компьютерных базах данных

Вероятно, вы уже пользовались электронными базами данных для поиска информации в школьной библиотеке. Наиболее известные из них: *Info Trac, Academic Abstracts* и *Medline.* Для психологов разнообразные поисковые средства предоставляет служба *PsycINFO* ассоциации *APA.* Некоторые из них выпущены в форме печатных изданий, но большинство доступно в электронном виде. Главная база данных, на основе которой строятся все остальные, называется *PsycINFO.* Она содержит ссылки на журнальные статьи, докторские диссертации, технические отчеты, книги и главы из книг. Самые старые ссылки датируются 1887 г. — годом основания первого журнала по психологии — «Американского психологического журнала». База охватывает литературу более чем 45 стран и включает ссылки на более чем 1300 журналов. Каждый месяц в базу данных добавляется около 5500 новых наименований. Для поиска самой свежей информации, зайдите на веб-сайт базы данных *PsycINFO* по адресу www.apa.org/psycinfo.

Название:	Влияние безошибочного упражнения на обучение детей дифференциации букв алфавита
Автор(ы):	Байрон Иглэнд
Место проведения	Университет Миннесоты, Программа обучения школьной психологии
Источник:	Журнал по прикладной психологии, № 60 (4), август, 1975 г., с. 533-536
Номер (ISSN/ISBN):	021-9010
Язык:	Английский
Краткое описание:	108 детей дошкольного возраста (4-5 лет) обучали дифференцировать буквы алфавита по: а) методике безошибочного упражнения, в которой в качестве очевидного стимула использовались выделенные отличительные особенности дифференцируемых букв; б) методике безошибочного упражнения, в которой отличительные особенности букв не выделялись, и в) традиционной методике затухающего подкрепления. Первая группа, занимавшаяся по методике безошибочного упражнения, во время обучения делала меньше ошибок, чем группа, обучавшаяся по методике затухающего подкрепления. Послеэкспериментальные исследования, одно, проведенное сразу по окончании эксперимента, и другое, отложенное, показали, что первая группа делала заметно меньше ошибок по сравнению с двумя другими группами - это говорит о том, что методика безошибочного упражнения эффективна, если отличительное особенности дифференцируемой буквы во время обучения выделяются. (© 1999 *APA/ PsycINFO.* Все права защищены.)
Ключевые фразы:	Безошибочное упражнение с выделением или без выделения отличительных особенностей или традиционная методика затухающего подкрепления; дифференциация букв алфавита; дети 4-5 лет
Темы:	Методы обучения; ошибки; буквы-алфавит; зрительная дифференциация; стимулы; характерные черты; дети дошкольного возраста
Классификация:	2300 Экспериментальная психология человека
Популяция:	Люди
Возрастная группа:	Дошкольный возраст (2-5 лет); детство (рождение - 12 лет)
Дата публикации:	19751001
База данных:	*PsycINF0*1967 - настоящее время

Рис. 3.6. Пример записи в базе данных *PsycINFO,* содержащей информацию об эксперименте Игленда (1975) по обучению детей дифференциации букв

Вы как студенты можете получить доступ к одному из двух видов базы данных *PsycINFO*. В вашей библиотеке может быть абонемент на пользование базой данных в виде компакт-дисков (обычно называемой *PsycLIT)*. Но, что более вероятно и удобно для вас, библиотека может иметь лицензию на пользование базой данных *PsycINFO* через сеть вашего кампуса с помощью службы *EBSCOhost*. Чтобы научиться пользоваться этой базой данных, лучше всего сесть за компьютер и поэкспериментировать, начав со страниц «Помощь». *APA* также выпускает различные руководства, которые вам поможет подобрать библиотекарь. Но лучше всего вам помогут опытные студенты психологических специальностей. Они смогут показать вам, как расширить или сузить поиск, и дадут советы о его более эффективной организации.

Когда в результате поиска вы получите список конкретных исследовательских статей, вы можете попросить *PsycINFO* показать их «записи», которые далее могут быть распечатаны или загружены на диск. На рис 3.6 показан вид стандартной записи *PsycINFO* для журнальной статьи (данная статья обсуждалась ранее в этой главе). Как видно на рисунке, каждая запись содержит несколько важных разделов. Эти разделы называются «полями» и кроме прочего включают название статьи, ее автора(ов), все необходимые выходные и прочие данные (название журнала, номер, страницы), краткий обзор статьи и дескрипторы (термины, которые можно использовать для дальнейшего поиска). Прочитав обзор, вы поймете, подходит ли вам данная статья, и если подходит, то сможете найти ее в коллекции журналов вашей библиотеки или заказать из другой.

Советы по поиску

Опыт — лучший учитель, но можно дать некоторые советы, которые помогут вам стать умелыми пользователями. Во-первых, если вы не уверены, с каких терминов начать поиск, воспользуйтесь встроенным в *PsycINFO* словарем *(Thesaurus)*. Чтобы открыть словарь, пользуйтесь иконкой «предметный поиск» *(subject search)*. Словарь — это электронная версия книги, изданной ассоциацией *APA* («Словарь психологических терминов», *Thesaurus of Psychological Index Terms),* предоставляющая расширенную систему перекрестных ссылок на термины, использующиеся как ключевые слова при поиске. Например, если попросить словарь найти тему «шизофрения», он выдаст ряд возможных терминов для поиска, «относящихся к» шизофрении (например, эмоциональная анестезия), «более широких, чем» шизофрения (например, психоз) и «более узких, чем» шизофрения (например, гебефрения). Если щелкнуть непосредственно на одном из этих терминов, то запустится поиск. Словарь также сообщит вам, когда термин был введен в базу данных (шизофрения — в 1967 г.) и что такие термины как «хроническая шизофрения» не являются дескрипторами для поиска в базе данных *PsycINFO* (для этого используется термин «шизофрения»). Если вы для поиска укажете этот термин, то вам сообщат, что «хроническая шизофрения» перестала быть поисковым термином в 1988 г. и попросят использовать термин «шизофения».

Кроме поиска по терминам база данных *PsycINFO* позволяет искать нужную информацию по именам авторов и по названиям журналов. Также можно искать

по «типу публикации» — это великолепный способ сконцентрироваться на особых типах статей или книг. К примеру, меня интересует история психологии, и я могу искать «некрологи». Также можно выбрать другие типы публикаций, например «лонгитюдные исследования», «повторение экспериментов», «обзор литературы» или «метаанализ».

Хорошая стратегия для поиска — это начать с последних лет и двигаться назад в прошлое. Кроме того, что вы при этом начнете с относительно новых исследований, такой способ поможет вам сузить поиск. Так, если вы пользуетесь термином «когнитивные карты», вы найдете более 100 статей, а если запрашиваете «когнитивные карты *и* статьи, опубликованные в 2001 г.», то обнаружите около десятка. Также в разделе «тип публикации» полезно искать термины «обзор литературы» или «метаанализ». Такие публикации обобщают результаты многих других статей. Ссылки, приведенные в записях к данным терминам, будут весьма ценны для ваших исследований. Обратите особое внимание на статьи в журнале «Психологический бюллетень» *{Psychological Bulletin)* и на главы из книжной серии «Ежегодный обзор психологии» *{Annual Review of Psychology).* В обоих изданиях печатают длинные обзоры литературы, которые являются потенциальными золотыми жилами, так как содержат расширенные списки ссылок. Как только вы найдете подходящую статью по вашей теме, вы можете использовать ее раздел ссылок как средство дальнейшего поиска. По этим ссылкам можно выбрать новые поисковые термины и выяснить имена исследователей, имеющих много публикаций по нужной теме.

Еще один совет — использовать функцию «и» для сужения поиска и функцию усечения, чтобы не сузить его слишком сильно. Предположим, к примеру, что вы занимаетесь социальной психологией и ищете статьи по исследованиям альтруизма. Кроме того, допустим, вы хотите рассмотреть альтруизм с точки зрения эволюционной психологии. Если вы просто запросите слово «альтруизм», вы получите слишком много записей. Проведенный мною для иллюстрации поиск выдал 1797 записей (к моменту чтения вами этой книги их будет гораздо больше). Однако когда я попросил *PsycINFO* искать слова «альтруизм И эволюционная психология», то было получено всего четыре записи. Так, конечно, гораздо легче справиться с материалом, но может оказаться, что есть и другие статьи на данную тему. В нашем случае это действительно так. Запросив термин «эволюционная психология», я «отрезал» записи, содержащие термины «эволюция» или «эволюционный». Чтобы решить эту проблему, термины, в данном случае относящиеся к «эволюции», должны быть «усечены» (т. е. укорочены). Для этого называют только несколько первых букв, а далее ставят звездочки. Например, запросив «эвол*», можно получить все термины, начинающиеся с этих четырех букв, в том числе эволюционная, эволюция и эволюционировать. На запрос «альтруизм и эвол*» я получил 178 записей. Это лучше, чем четыре, но слишком много, чтобы начать просматривать их по очереди. Следующий шаг — начать с наиболее поздних статей. Я указываю «альтруизм и эвол*» за 1995-1999 гг. и получаю 37 записей. Это подходящее количество для просмотра записей с краткими обзорами.

Приобретя опыт работы с базой данных *PsycINFO,* вы научитесь распознавать полезную информацию на интересующую вас тему. Начав изучать эту информа-

цию и обсуждать ее с другими студентами и преподавателями, вы постепенно приобретете знание литературы (об эффективном чтении исследовательских статей см. табл. 3.1). Знание литературы, в свою очередь, благоприятно скажется на разработке вами идей для исследований и их формулировке в виде эмпирических вопросов. Поставив правильные эмпирические вопросы, вы сможете разрабатывать эксперименты, отвечающие на них. Тема разработки исследований будет вскоре нами рассмотрена, но сначала необходимо изучить темы, связанные с данными, собираемыми для ответа на эмпирический вопрос.

Таблица 3.1

Как получить максимальное количество информации из журнальной статьи

В какой-то момент, в ходе курса по методам исследований (возможно, как часть задания по написанию обзора литературы) вам придется читать статьи об исследованиях, опубликованные в одном из психологических журналов. Важно помнить, что журнальные статьи написаны не для студенческой аудитории, — они нацелены на профессиональных исследователей. Поэтому их может быть очень трудно читать. Ниже представлены некоторые советы, призванные облегчить эту задачу.

- Как можно внимательнее прочитайте краткое изложение. Оно представляет собой общее описание исследования и, возможно, является самым простым для чтения и понимания разделом, поэтому прочитайте его несколько раз.
- В одном или двух первых абзацах найдите общую формулировку исследуемой проблемы. Примите к сведению, что нужный раздел не будет озаглавлен «Введение», но будет включать весь текст между кратким изложением и «Методами».
- В конце введения, скорее всего в последнем абзаце, найдите развернутую формулировку гипотезы или гипотез, проверяемых в эксперименте. Эти гипотезы основываются на формулировке проблемы и исследовательских вопросах, поставленных в экспериментах, описанных в средней части введения. Запишите гипотезы и всегда помните о них во время чтения.
- При чтении раздела «Методы» особое внимание уделите описанию процедуры исследования и экспериментальному плану. Представьте себя на месте участников и сформируйте четкие представления о том, что они должны делать в ходе исследования. Если это экспериментальное исследование, выпишите зависимые и независимые переменные (вы уже знакомы с этими терминами из курса общей психологии и узнаете о них более подробно в главе 5).
- Раздел «Результаты» может оказаться особенно трудным, так как он содержит статистическую информацию и обозначения, которые могут оказаться вам незнакомыми. Хорошо написанный раздел «Результаты» должен содержать ясное словесное описание полученных результатов, а также графики и таблицы, облегчающие понимание. Он должен дать возможность понять суть происходящего в ходе исследования без привлечения чисел или условных обозначений.
- Последняя часть статьи — это раздел «Обсуждение». Он часто начинается с общего описания результатов, так что если раздел «Результаты» остался вам непонятен, еще не все потеряно. Главная задача этого раздела — объяснить результаты с точки зрения исходных гипотез, поэтому автор будет выстраивать параллели с разделом «Введение». Последнее, на что необходимо обратить внимание в данном разделе, — на описание исследования, которое должно быть проведено вслед за данным (вопрос «что дальше?»). Раздел «Обсуждение» является прекрасным источником идей для новых исследований.

Резюме

Разновидности психологических исследований

Фундаментальные исследования по психологии нацелены на открытие базовых принципов поведения, а прикладные исследования проводятся для решения практических проблем. И фундаментальные, и прикладные исследования могут проходить как в лаборатории, так и в полевых условиях. Лабораторные исследования обеспечивают более полный контроль за ходом экспериментов, а полевые — лучше воспроизводят ситуации реальной жизни. Исследования, вовлекающие участников в процедуру (т. е. создающие экспериментальную реальность), даже если при этом создаются ситуации, далекие от условий обычной жизни, могут стать источником важной информации о поведении. Большинство исследований в психологии — количественные по своей природе и включают статистический анализ численных данных, но в последние годы наблюдается рост числа качественных исследований (таких как содержательный анализ формализованных интервью).

Постановка эмпирических вопросов

Начальный этап любого исследовательского проекта — формулировка эмпирического вопроса, т. е. такого вопроса, на который можно ответить с помощью объективных данных. Эмпирические вопросы содержат термины, определенные достаточно точно, чтобы обеспечить возможность воспроизведения исследования (операционально). Некоторые исследования, посвященные одной теме, могут иметь разные операциональные определения одних терминов и при этом приводить к одинаковым выводам (например, о том, что степень голода влияет на скорость обучения прохождению лабиринта).

Разработка исследований на основе наблюдения

Идеи некоторых исследований возникают в ходе размышлений над замеченным в ежедневной жизни, особенно над событиями, достаточно необычными, чтобы привлечь к себе внимание. Конкретные проблемы, которые необходимо решить, также приводят к разработке исследования — многие прикладные исследования в целом и исследования по оценке программ в частности разрабатываются именно таким способом. Иногда мы наблюдаем события, которые происходят неожиданно или случайно. Случайные наблюдения часто приводят к возникновению идей дальнейших исследований (например, исследование, обнаружившее «эффект края»).

Разработка исследований на основе теорий

Теории обобщают и организуют существующие знания, являются основой для прогнозирования и дают рабочее объяснение явлениям. Между построением теории и исследованием существует двустороняя связь. Эмпирические вопросы дедуктивным методом выводятся из теории и воплощаются в эксперименте. Сделанные на основании экспериментов выводы либо поддерживают, либо опровергают теорию. Нельзя доказать истинность теории, но можно ее опровергнуть, хотя бы в принципе. Однако в действительности теория отвергается только после того, как ученые многократно убедятся в том, что на ее основе невозможно делать предска-

зания. В большинстве случаев теории развиваются с учетом накопленных знаний о каком-либо явлении. В психологии теории приносят пользу если они служат исследованиям, углубляющим наше понимание поведения. Хорошим теориям присуща простота и точность формулировки, позволяющая подвергнуть их фальсификации с помощью правильно разработанного исследования.

Разработка исследований на основе других исследований

В психологии исследователи редко думают в терминах изолированных опытов. Напротив, они разрабатывают программу исследований, состоящую из набора взаимосвязанных экспериментов, касающихся определенной научной области. Исследователи постоянно используют результаты экспериментов для формирования дальнейших исследований. Исследовательские программы нередко включают повторение существующих открытий, а также их дополнение и распространение на другие области исследований.

Креативное мышление в науке

Научная креативность работает, когда исследователи устанавливают связи между идеями или событиями, которые большинство людей считают не связанными между собой. Креативные ученые обладают глубокими познаниями в определенной области и замечают важность событий, на первый взгляд не имеющих отношения к актуальной проблеме.

Обзор литературы

Эмпирические вопросы чаще приходят в голову исследователям, знакомым с литературой, посвященной исследованиям в определенной области. В основном знакомство с ней устанавливается с помощью поиска в электронных базах данных, таких как *PsycINFO*.

Задания для повторения

Выбор ответа

1. Какие из следующих утверждений, сравнивающих полевые исследования с лабораторными, верны?
   1) лабораторные исследования моделируют более жизнеподобные ситуации, т. е. проявляют больший жизненный реализм;
   2) для достижения экспериментального реализма исследования должны проводиться в лаборатории;
   3) лабораторные исследования обеспечивают более высокий уровень контроля за условиями проведения экспериментов;
   4) при проведении лабораторных исследований обоснованное согласие получить сложнее, чем при проведении полевых исследовний.
2. Каковы преимущества использования операциональных определений?
   1) они вынуждают исследователей разных лабораторий использовать абсолютно одинаковые определения для данного психологического конструкта;

2) они облегчают процесс воспроизведения исследования;
3) они повышают жизненный реализм исследования;
4) все перечисленное выше неверно — из-за отсутствия общепринятых определений конструктов операциональные определения используются редко.

3. Одно из предсказаний на основе теории выученной беспомощности: если животные имели предварительный положительный опыт избегания воздействия, то вероятность формирования у них стандартной беспомощности в ходе неконтролируемого воздействия электрического шока будет ниже. Представьте, что вы провели исследование для проверки этой идеи и оно показало, что у собак, в начале эксперимента успешно избегавших шока, в дальнейшем не вырабатывается беспомощность. Какой вывод можно сделать на основании этих фактов?
   1) теория получила дедуктивное подтверждение;
   2) доказана истинность теории;
   3) теория получила индуктивное подтверждение;
   4) подтверждений теории получено не было.
4. В исследовании 1 проводилось сравнение восприятия 2- и 4-месячных детей. В исследовании 2 использовались те же задания, но сравнивались четыре группы детей в возрасте 2, 4, 6 и 8 месяцев. Второе исследование:
   1) имело более высокий уровень экспериментального реализма, чем первое;
   2) является примером сходящейся операции;
   3) является хорошим примером повторения и дополнения;
   4) является избыточным — нет необходимости включать первые две группы (детей в возрасте 2 и 4 месяцев).
5. Выражение «случай благосклонен к подготовленному уму» говорит о том, что:
   1) креативные ученые имеют глубокие познания в своей области;
   2) креативность — это дело удачи (случая);
   3) наиболее креативны те, кто незнаком с исследуемой темой;
   4) случайные события помогают только наиболее подготовленным исследователям.

Короткие эссе

1. Почему эксперименты Игленда по обучению чтению являются примером прикладного исследования, а эксперименты Найссера — фундаментального?
2. Каковы сравнительные достоинства и недостатки исследований, проводящихся в пределах и за пределами лаборатории?
3. Почему в исследовании Даттона и Эрона по изучению влюбленности ученые, имея результаты полевого исследования, сочли необходимым провести повторные лабораторные эксперименты?

4. Приведите три различных операциональных определения голода и объясните, почему исследование, использующее все три определения, может быть примером сходящейся операции.
5. Что такое теория в психологии и каковы особенности хороших теорий?
6. На примере теории выученной беспомощности проиллюстрируйте двусторонние отношения между теорией и фактами. При ответе используйте термины: индукция, дедукция и гипотеза.
7. Объясните, почему ученые обычно не говорят: была *доказана* истинность теории.
8. Объясните, как эксперимент с Умным Гансом иллюстрирует важность а) стратегии фальсификиции и б) использования наиболее простых объяснений.
9. Что такое предварительное исследование и с какой целью оно проводится?
10. На примере идеи применения лабиринта к изучению поведения проиллюстрируйте роль креативного мышления в науке.

Упражнения

Упражнение 3.1. Что дальше?

Вспомните фундаментальное открытие, сделанное Милгрэмом при исследовании подчинения. Взрослые мужчины полностью подчинялись приказу применить электрошок к «обучающимся» в наказание за ошибки. Более половины участников повысили силу тока до 450 вольт. Представьте, что вы входите в исследовательскую группу Милгрэма и первое исследование завершено.

1. Имея результаты первого исследования, как вы построили бы следующее?
2. Как вы думаете, что произойдет (т. е. какова ваша гипотеза)?

Придумайте не меньше трех различных вариантов процедуры исследования, которая увеличит наши знания о факторах, обусловливающих подчинение.

Упражнение 3.2. Выработка операциональных определений

Выработайте два различных операциональных определения для каждого из следующих психологических конструктов:

1) фрустрация;
2) застенчивость;
3) тревога;
4) чувство направления.

Упражнение 3.3. Искажение по подтверждению

В этой главе мы узнали о научной стратегии, заключающейся в получении эмпирических выводов посредством исключения или *фальсификации* альтернативных объяснений. Как показывает упражнение Уэйсона и Джонсон-Лэйрда (Wason and Johnson-Laird, 1972, p. 172-173), эту стратегию довольно трудно реализовать. Проверьте это.

Представьте, что вы держите в руках четыре карточки, на каждой из которых с одной стороны напечатана буква, а с другой — цифра. Посмотрев на карточки, вы видите следующее:

E K 4 7

Необходимо решить, какую карточку следует перевернуть, чтобы проверить истинность следующего правила:

Если на одной стороне карточки изображен гласный звук, то на другой стороне будет изображено четное число.

Какую карточку вы перевернете? (Подсказка: думайте в терминах фальсификации: какая карточка, будучи перевернутой, может фальсифицировать высказывание?)

Упражнение 3.4. Поиск в базе данных *PsycINFO*

В базе данных *PsycINFO* найдите записи о любых пяти статьях, упомянутых в этой главе. Для каждой из пяти статей а) найдите другую статью того же автора, б) найдите другую статью на ту же тему, напечатанную в последние три года.

ГЛАВА 4

Измерения, выборка и обработка данных

Обзор задач главы

В этой главе мы начнем подробное рассмотрение процедуры сбора данных в ходе исследования. Мы будем говорить о разнообразных параметрах, измеряемых в психологических исследованиях, о факторах, определяющих, имеют ли смысл проводимые измерения; рассмотрим различные типы так называемой шкалы измерений и методы отбора участников исследования. В этой главе также даются начальные сведения о различиях между описательной статистикой и статистикой вывода (некоторые из вас, я надеюсь, смогут пополнить уже имеющиеся знания) и рассматривается логика проверки гипотез. Изучив данную главу, вы сможете:

- понять взаимоотношения между конструктом (например, зрительным образом) и измеряемым параметром действия (например, скоростью реакции);
- различать надежность и различные виды валидности измерений;
- описать особенности номинальной, порядковой, интервальной шкал и шкалы отношений, а также указать, когда каждая из них должна использоваться;
- описать три вида вероятностной выборки (простую случайную, расслоенную и кластерную), а также указать, когда каждая из них должна использоваться;
- понять, когда нужно использовать вероятностную выборку и почему в большинстве психологических исследований используются простые (невероятностные) выборки;
- обобщать данные, определяя общую тенденцию (например, вычисляя среднее арифметическое), изменчивость (например, находя стандартное отклонение) и делая визуальные представления данных (например, методом стебля и листа);
- понимать суть процедуры проверки гипотезы и знать составляющие процесса заключительной обработки данных.

Из главы 1 вы узнали, что психологи-исследователи применяют информационный подход: выводы о поведении должны основываться на данных, собранных научными методами. Решить, параметры каких именно действий измерять, как прово-

дить измерения и как организовать полученный набор чисел, весьма нелегко. Данная глава открывает обсуждение взаимоотношений между собираемыми данными и психологическим знанием, которое мы продолжим в следующих главах.

Что измерять: разнообразие поведенческих актов

Разнообразие поведенческих актов, параметры которых измеряются психологами-экспериментаторами, практически не ограничено. Поддающиеся измерению феномены варьируются от внешних форм поведения, как передвижение по лабиринту, до получаемых с помощью опросников вербальных отчетов и записей физиологической активности, проявляемой при выполнении некоторых заданий. Для иллюстрации многообразия измерений в ходе психологических исследований рассмотрим следующие примеры:

1. В работе «Зависимость объема восприятия больных шизофренией от отвлекающих воздействий и латеральности» (Elkins, Cromwell & Asarnow, 1992) изучались ограничения объема восприятия у больных шизофренией. В сравнении с контрольной группой эти испытуемые хуже справились с заданием на распознание отдельных букв в массиве других, отвлекающих внимание. Оценивалось, могут ли испытуемые точно указать нужные буквы в различных ситуациях.
2. В работе «Способности к математическим вычислениям у детей младшего возраста из семей со средним и низким достатком» (Jordan, Huttenlocher & Levine, 1992) исследовалась способность детей к решению различных математических задач. Было обнаружено, что дети из семей со средним достатком лучше выполняли задания, если их формулировали устно, при этом обе группы одинаково успешно справлялись с невербальными задачами. Оценивались две характеристики: ответ на вопрос задачи и используемая стратегия. Детей наблюдали во время выполнения задания, следя за тем, используют ли они счет на пальцах.
3. В исследовании «Восприятие социальной поддержки, социальные навыки и особенности отношений с окружающими у женщин, страдающих булимией» (Grissett & Norvell, 1992) женщины с булимией и контрольная группа выполняли батарею тестов самоотчета. Оценивались ответы на поставленные вопросы (например, по шкале восприятия поддержки). Кроме того, с участницами проводились краткие беседы, записываемые ассистентами и оцениваемые наблюдателями. Речевое и неречевое поведение женщин в процессе беседы оценивалось по показателю «социальная эффективность». Женщины с булимией считали, что не получают большой социальной поддержки, их социальная эффективность была невысока.
4. В работе «Определяется ли выбор направления в радиальном лабиринте с помощью когнитивной карты?» (Brown, 1992) на этот вопрос был получен ответ: «Не всегда». Радиальный лабиринт, состоящий из центральной площадки и дорожек, расходящихся от нее во всех направлениях, — один из наиболее популярных видов современных лабиринтов. Его часто используют при изучении пространственной памяти (когнитивных карт) у крыс. В ис-

следовании Брауна изучалась способность крыс не пользоваться повторно одними и теми же дорожками. Оценивались ситуации «макровыбора», операционально определенные как возникающие, когда крыса заглядывает за поворот дорожки, и «микровыбора», обозначенные как ситуации, когда крыса поворачивается к дорожке, но не идет по ней (все это фиксировалось с помощью видеокамеры, направленной вниз на лабиринт).

Разработка измерений на основе конструктов

Из приведенных примеров видно, что исследователи оценивают параметры поведения различными способами. Но как они решают, что необходимо измерять; что можно определять вычислительные стратегии, наблюдая счет на пальцах, фиксировать микровыборы, наблюдая положение тела, а степень внимания — через выбор букв из массива?

Ученые знают что надо измерять, в частности, потому, что они знакомы с литературой по теме своего исследования и знают, какие измерения в этих случаях обычно проводятся. Новые измерения возникают также из модификации общеизвестных в психологии вообще. Кроме того, разработка измерений может проводиться в процессе усовершенствования исследуемых конструктов. Рассмотрим эти вопросы более подробно.

При планировании исследования одна из основных задач — как можно более точно определить конструкты, используемые в проекте. Звучит знакомо? Так и должно быть, ведь мы опять говорим об операциональных определениях. В любой исследовательский план входит установление границ интересующих ученого конструктов, которые по определению не наблюдаются непосредственно, и выявление типов поведения, адекватно отражающих эти конструкты. В приведенных выше примерах все исследователи ставили перед собой одну и ту же задачу: на основе определенного явления построить управляемый эксперимент, определив конструкты в терминах измеряемых характеристик поведения. В табл. 4.1 представлена классификация четырех исследований по изучаемым конструктам и способам их операционального определения на основе конкретных видов поведения.

Таблица 4.1

Примеры конструктов и способы их измерений

Конструкт	Параметр **поведения, на основании которого измерялся конструкт**
Объем внимания	Точность узнавания букв
Способность к вычислениям	Количество правильно решенных задач
Воспринимаемая социальная поддержка	Результаты теста самоотчета
Социальная эффективность	Оценка наблюдателем социальных взаимодействий
Макровыбор у крыс	Заходит или нет крыса на дорожку, ведущую от центра в радиальном лабиринте
Микровыбор у крыс	Ориентация тела крысы в направлении дорожки, ведущей от центра в радиальном лабиринте

Легко заметить, что ни один из этих конструктов (внимание, способность к вычислениям) нельзя наблюдать непосредственно — выводы о них делаются на основании проведенных измерений. Психологи обращаются к этому методу снова и снова, что позволяет исследователям ставить такие эмпирические вопросы, ответы на которые на первый взгляд невозможно получить. Рассмотрим более подробно пример двух процедур, которые часто используются при изучении вопросов, получение эмпирических ответов на которые может представляться трудным или даже невозможным.

Понимают ли младенцы, что существует сила тяжести?

Можно ли продемонстрировать, что люди используют зрительные образы в процессе обработки информации?

Измерения, необходимые для изучения этих, на первый взгляд неэмпирических вопросов, так же просты, как в случае исследований вопросов: а) как долго младенец может смотреть на один объект или б) сколько времени требуется людям, чтобы принять решение.

Пример 1. Привыкание

Знают ли младенцы о существовании силы тяжести? Как вообще возможно это определить? Конечно, их нельзя спросить прямо, но вопрос можно задать опосредованно, с помощью методики измерения количества времени, которое затрачивает ребенок, глядя на различные объекты. Эта так называемая процедура привыкания заключается в том, что ребенку много раз показывают один и те же стимульный материал, а затем переключаются на новый. Из исследований известно, что младенцы любят новизну (Spelke, 1985), так что если им многократно демонстрируется один и тот же объект, они теряют к нему интерес (т. е. прекращают смотреть на него). Термин «привыкание» определяется как постепенное ослабление ответной реакции на повторяющиеся стимулы. Когда показывают новый объект и младенец воспринимает его как нечто новое или необычное, то он будет смотреть на него дольше. Таким образом, если время разглядывания объекта уменьшается, а затем внезапно увеличивается, то можно предположить, что младенец заметил что-то новое.

Имея это в виду, рассмотрим восхитительный эксперимент, проведенный Кимом и Спилке в 1992 г. Они сравнили 5- и 7-месячных детей и выяснили, что представление о силе тяжести формируется именно в этот период жизни. Сначала детям многократно показывали кадры, на которых шары катились вверх или вниз по наклонным плоскостям, как показано на верхних двух изображениях на рис. 4.1. Во время процедуры привыкания одним младенцам показывали шар, катящийся с ускорением вниз по плоскости (условие 1), а другим — шар, катящийся вверх по плоскости, замедляющий свое движение (условие 2). Эти явления отражают естественное воздействие силы тяжести на шары, катящиеся вверх или вниз по склону. Когда привыкание было сформировано (т. е. время рассматривания после повторных попыток заметно снизилось), детям показали либо «естественное тестовое явление» (изображения в середине), либо «неестественное тестовое явление» (нижние изображения). Обратите внимание, что естественное тестовое явление отличается от явления, формирующего привыкание, в двух аспектах: в аспектах направления и скорости. При этом если младенец при формировании привыкания

наблюдал движение с ускорением вниз, то ему показывали естественное явление с шаром, катящимся вверх и замедляющим свое движение. Неестественное явление отличается только одним показателем — направлением. Младенцу, во время привыкания наблюдавшему ускоряющееся движение шара вниз, показывали неестественное явление с шаром, катящимся в другую сторону, но с такой же скоростью (т. е. с ускорением). Исходя из того в чем тестовое явление отличаются от привычного, разумно предположить, что дети воспримут естественное явление как новое (изменилось два фактора) и будут дольше смотреть на него, чем на неестественное явление (изменился один фактор). И действительно, 5-месячные дети так и поступили. Однако 7-месячные дети дольше смотрели на *неестественное* явление. Вероятно, это произошло потому, что в случае неестественного явления нарушается закон тяготения, тогда как естественное явление следует закону, действие которого было показано на записи, использованной для привыкания. Таким образом, младшие дети обратили внимание на изменения в общем числе стимульных измерений, а старшие отмечали изменения, нарушающие закон силы тяжести. Измерив время, которое младенцы затрачивали на просмотр фильма, Ким и Спилке пришли к выводу, что младенцы в возрасте 7 месяцев обладают представлением о силе тяжести.

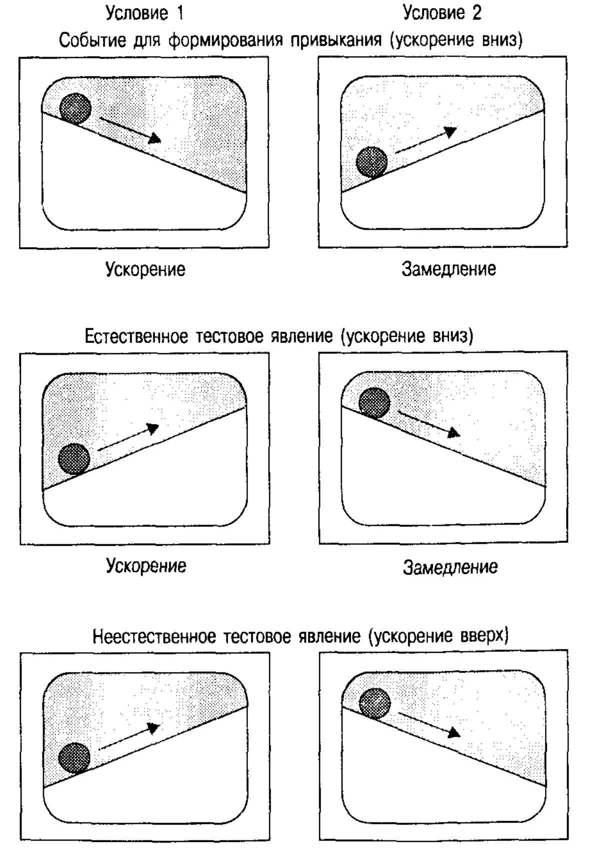


Рис. 4.1. Стимульные материалы из исследования с использованием привыкания, проведенного Кимом и Спилке в 1992 г.

Пример 2. Скорость реакции

Используем ли мы зрительные образы в процессе обработки информации? Как это можно выяснить? Конечно, об этом можно спросить, но если человек ответит «Да, я использую образы», то как узнать что именно он делает, используя их? Таким образом, перед вами встанет проблема, которая привела к отказу от интроспекции как метода, — отсутствие объективности. Однако можно попросить людей выполнить определенное задание, решая которое они будут вести себя одним образом, используя образы, и другим — не используя. Именно такую стратегию использовали Шепард и его коллеги в серии хорошо известных экспериментов по изучению так называемого «мысленного вращения».

Рассмотрите две пары геометрических фигур на рис. 4.2. Являются ли правые фигуры в каждой паре теми же самыми, что слева, только перевернутыми, или это другие фигуры? Как вы это определили? В 1971 г. Шепард и Мецлер в ходе эксперимента попросили участников не просто ответить на этот вопрос, но также зафиксировать *время,* за которое был найден ответ. Их идея заключалась в том, что если участники решают задачу, мысленно переворачивая левую фигуру так, чтобы она совпала с правой, то этот процесс займет определенное количество времени. Более того, чем сильнее должен быть поворот, чтобы фигуры совпали, тем больше времени на это уйдет. Я полагаю, вы понимаете, как это происходит. Шепард и Мецлер последовательно изменяли градус поворота и обнаружили, что с его увеличением растет время, необходимое для ответа на вопрос. Таким образом, измерив скорость реакции, они заключили, что участники используют мысленные образы.

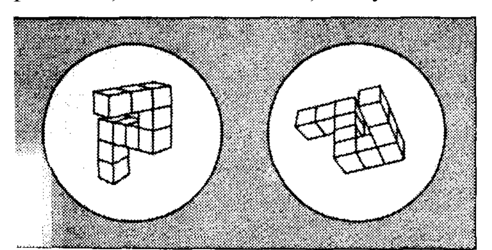
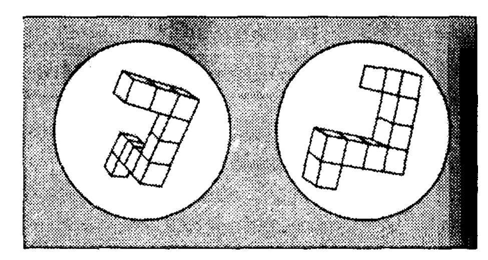



Рис. 4.2. Стимульные материалы для исследования мысленного вращения, проведенного Шепардом и Мецлером в 1971 г.

Измерение скорости реакции — один из старейших и наиболее устойчивых методов в психологии, но причины, по которым к нему прибегают, с течением времени изменяются. Дополнительные сведения о происхождении и развитии этого весьма эффективного метода, применяемого в экспериментальной психологии, вы получите, прочитав вставку 4.1.

Оценочные измерения

Сак можно определить правильность измерения поведения? Чем объясняется уверенность ученых в оценке таких вещей, как время наблюдения события, скорость реакции, выполнение теста интеллекта, определение восприятия социальной поддержки и т. д.? Чтобы ответить на этот вопрос, необходимо обсудить два ключевых фактора: надежность и валидность.

ВСТАВКА 4.1

История: скорость реакции — от хронометрических экспериментов до мысленного вращения

В психологической литературе понятие скорости реакции впервые упоминается в работе Ф. К. Дондерса, датского физиолога, который установил, что продолжительность психического явления можно определить, найдя разницу между скоростями реакций, получаемых при выполнении различных видов заданий (Boring, 1950). Его идея породила водопад исследований, посвященных явлению, сегодня известному как «психическая хронометрия». Сначала исследователи измеряли скорость простой сенсомоторной реакции (ПСР) - скорость однократной реакции, полученной сразу за предъявлением однократного стимула, например сигнала красного цвета. Далее они усложняли задание, показывая поочередно два стимула и предлагая испытуемому отвечать только на один из них. Это получило название скорости реакции выбора (РВ), так как испытуемому требовалось сначала дифференцировать стимулы, например красный и зеленый сигналы, а затем отвечать на один из них. СДР включает СПР плюс время психического явления «дифференциации», поэтому было решено, что найдя разность СДР и СПР можно получить продолжительность этого психического явления:

РВ = ПСР + дифференциация;

дифференциация = РВ - ПСР.

Процедуру можно разрабатывать и дальше, получая более дробные разности психических явлений - этот процесс очень воодушевлял исследователей конца XIX в., так как в то время они пытались заявить психологию как науку, а что может быть более научным, чем измерение психических явлений с точностью до долей секунд? Особенно высокой популярностью эта процедура пользовалась в лаборатории Вундта в Лейпциге, и как видно из рис. 4.3, она была быстро импортирована в Соединенные Штаты. На рисунке показан ход эксперимента по измерению скорости реакции, проводящийся в Университете Кларк в начале 1890-х гг.

К сожалению, вскоре обнаружились серьезные проблемы. В частности, в одних случаях скорость реакции была ниже, чем предсказывалось, а в других выше. Освальд Кюльпе, один из студентов Вундта, отметил главнейшую ошибку метода - более сложные психические явления не образуются из простых в результате их простого суммирования. Напротив, сложные психические явления обладают присущими только им качествами, которые не являются суммой более простых.

Хотя доводы Кюльпе быстро положили конец психической хронометрии и дискредитировали измерение скорости реакции как метод еще в период расцвета бихевиоризма (приблизительно 1930-1950 гг.), впоследствии она пережила возрождение в некоторых областях когнитивной психологии. Возникла идея применять психическую хронометрию не для измерения точной продолжительности психических явлений, а для проверки предположений, сделанных на основе когнитивных теорий. Хороший пример тому - феномен мысленного вращения. Шепард предположил, что если участники применяют мысленное вращение, то эта умственная активность должна занимать определенное количество времени (Shepard & Metzler, 1971). Поворот на больший угол должен занимать больше времени, и как вы могли убедиться, это действительно так.

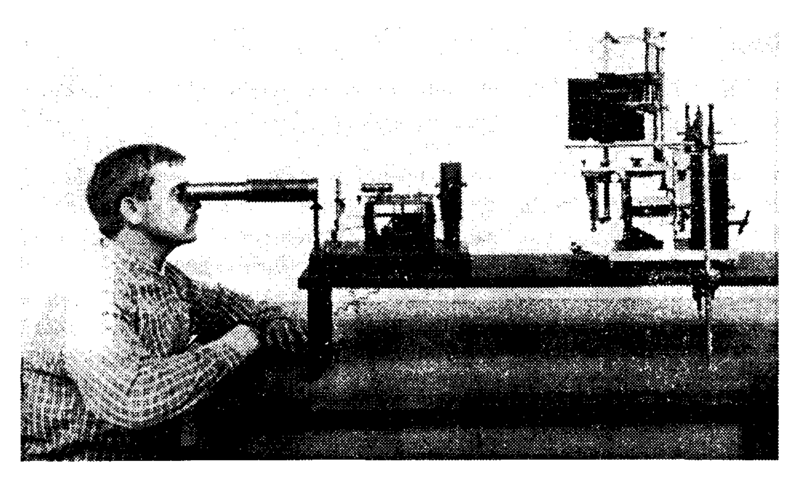


Рис. 4.3. Эксперимент по изучению скорости реакции, проводившийся в Университете Кларк около 1892 г. Ответная реакция выражалась в том, что участник, видя в трубе предъявляемый стимул, правой рукой переставал нажимать на телеграфный ключ

Надежность

В целом измерение характеристик поведения считается надежным, если его результаты повторяются при повторном измерении. Скорость реакции является хорошим примером — высокая надежность была одной из причин многолетней популярности таких измерений. Человек, при первой попытке реагирующий на красный сигнал через 0,18 секунды, наверняка покажет ту же скорость при следующих попытках, и практически все попытки будут давать результаты, близкие к 0,18 секунды. К примеру, оценки теста *GRE {Graduate Record Exam)* также относительно надежны. Человек, получивший 850 за общий тест *GRE,* во второй раз, скорее всего, получит сходную оценку, и очень маловероятно, что его балл будет 1350.

Из этих двух примеров можно понять, почему надежность измерений настолько важна. Без нее невозможно было бы определить, что означают оценки, полученные при конкретном измерении. Предположим, измеряя скорость реакции, вы хотите узнать, насколько быстро реагирует конкретный человек. Если это время станет сильно изменяться от попытки к попытке, то на данный вопрос ответить будет невозможно. Аналогично, если разница в оценках теста *GRE* для одного человека может составлять 400 или 500 баллов, вузы не смогут использовать результаты этого теста, так как по ним нельзя будет предположить действительный уровень знаний студента.

Надежность измерения характеристик поведения — это функция от количества **ошибок измерения.** Если ошибок много — надежность низкая, и наоборот. Ни одно подобное измерение не является абсолютно надежным, так как все они содержат некоторое количество ошибок. Это означает, что результат измерения представляет собой гипотетическую истинную оценку плюс ошибку измерения. В идеале ошибка измерения настолько мала, что полученная оценка близка к истинной.

Процедура измерения скорости реакции хорошо иллюстрирует, что такое ошибка измерения и как она влияет на надежность. Как и в предыдущем примере, предположим, что скорость реакции человека составила 0,18 секунды. Является ли это истинным значением измеренной скорости? Нет, как несложно заключить из следующих результатов, полученных при измерении скорости реакции у того же человека:

0,16с, 0,15 с, 0,19 с, 0,17 с, 0,19 с.

Результаты различаются, потому что каждая из попыток не обошлась без ошибки измерения. Эта ошибка вызвана различными факторами — некоторые из них действуют случайным образом от попытки к попытке. Например, при одной из попыток человек может реагировать быстрее, чем предполагает истинное значение, если будет ожидать появления стимула, или медленнее из-за кратковременного снижения внимания. Также может возникнуть систематическая ошибка, если, к примеру, перед тем, как предъявить стимул, экспериментатор дает участнику сигнал приготовиться, а время между этим сигналом и стимулом остается постоянным. В этом случае участник начнет предугадывать стимул и скорость реакции будет систематически выше, чем истинная оценка.

Несмотря на небольшую ошибку измерений, приведенные выше оценки неплохо соответствуют друг другу и измерение скорости реакции в этом случае будет считаться более надежным, чем при следующем наборе оценок, полученном вслед за результатом 0,18 секунды:

0,11с, 0,25 с, 0,19 с, 0,09 с, 0,31с.

При результатах, варьирующихся от одной десятой секунды до одной третьей, очень трудно установить действительную скорость реакции человека.

Если оценки надежны, то исследователь может в зависимости от их значений приписать им некоторый смысл. Надежность также позволяет исследователю делать более осмысленные сравнения одних оценок с другими. Например, сравнение первого набора оценок (0,16 с, 0,15 с и т. д.) с нижеследующим обнаруживает разницу в общей скорости реакции:

0,23 с, 0,26 с, 0,21с, 0,22 с, 0,24 с.

Очевидно, что второй испытуемый реагировал медленнее, чем первый.

Есть разные способы вычисления надежности, но в экспериментальных исследованиях это редко проделывают. Уверенность в надежности измерений вырабатывается со временем в результате процесса их повторения. Например, процедура вырабатывания привыкания и измерение скорости реакции проводились достаточно часто и приносили весьма непротиворечивые результаты, так что исследователи убедились в их надежности.

Надежность вычисляется более строго в исследованиях, посвященных оценке адекватности всевозможных психологических тестов, разработанных для измерения таких конструктов, как личностные факторы, способности (например, *IQ)* и установки. Это обычно письменные тесты, в которых требуется отвечать на вопросы или что-либо утверждать/опровергать. В описанном выше исследовании були-

мии участницы заполняли различные тесты, в том числе так называемую шкалу восприятия поддержки. Анализ, призванный установить надежность такого вида тестов, требует применения корреляционных процедур. Например, участники могут выполнить тест в двух различных ситуациях, а далее будет определена степень сходства двух наборов оценок. Если не повлияют серьезные изменения, произошедшие в жизни участников, результаты двух измерений по шкале восприятия поддержки должны быть похожи. Степень сходства выражается в терминах корреляции (высокое сходство = высокая корреляция). В главе 9 будут более подробно объяснены особенности этого вида анализа и его значение для психологического тестирования в целом.

Валидность

Измерение характеристик поведения считается валидным, если измеряют именно то, что собирались. В ходе оценки восприятия социальной поддержки должно действительно измеряться количество поддержки, которую (по их мнению) получают испытуемые, а не какой-либо другой конструкт. Тест на интеллект должен измерять интеллект, а не что-то иное.

Иногда считается, что измерение обладает определенной валидностью просто потому, что оно представляется разумным. При измерении интеллекта тест с задачами, требующими размышлений, представляется более подходящим, чем тест, в ходе которого участники должны аккуратно проехать на велосипеде между двумя белыми линиями. Это значит, что решение задач и рассуждения имеют большую очевидную валидность для измерения интеллекта, чем езда на велосипеде (которая обладает очевидной валидностью для определения способности к удержанию равновесия). Но конечно, только очевидной валидности недостаточно — тест может выглядеть весьма разумным, но при этом не быть валидным. Большинство опросов, которые публикуются в популярных журналах, попадают именно в эту категорию.

Более строгий вид валидности называется критериальной валидностью. Она показывает: а) можно ли на основании измерения предсказать будущее поведение и б) связано ли это измерение с другими измерениями поведения. Например, чтобы тест на уровень интеллекта был валидным: а) должно быть возможным на его основании предсказать успеваемость ребенка в школе и б) его результаты должны быть сходны с результатами других известных тестов по измерению уровня интеллекта. Валидность названа «критериальной», потому что результаты рассматриваемых измерений сопоставляются с некоторым значением или критерием. В данном примере критериальными переменными являются: а) будущие оценки в школе и б) оценки по существующим тестам на интеллект. Так же как оценка надежности, исследование критериальной валидности корреляционное по своей природе и проводится в первую очередь для оценки психологических тестов. В главе 9 вы снова встретитесь с этим вопросом.

Еще один вид валидности — конструктная валидность — особенно подходит для экспериментальных исследований. Она имеет отношение сразу к двум вопросам: является ли оцениваемый конструкт валидным, а инструмент, с помощью которого производится оценка, — наилучшим. Конструктная валидность тесно свя-

зана с сущностью теории, построением гипотезы на основании теории и оценкой теорий, выведенных из результатов исследования. Конструктная валидность как таковая никогда не подтверждается и не разрушается одним исследованием, а также ее невозможно доказать по тем же причинам, но каким невозможно доказать теорию. Уверенность в конструктной валидности растет постепенно по мере получения исследованиями подтверждающих результатов.

Как мы видели, аппарат психологии насыщен конструктами, которые не могут наблюдаться непосредственно, — такими понятиями, как голод, тревожность, интеллект, депрессия, воспринимаемая социальная поддержка, объем внимания и т. д. Так как эти конструкты являются гипотетическими, их наличие может быть выведено только из специально проведенных измерений. Исследование, в котором проводится измерение конструкта определенным образом и которое дает ранее предсказанные результаты, подтверждает не только сам конструкт, но и инструмент, использованный для его измерения.

Пример 3. Конструктная валидность

Для примера конструктной валидности рассмотрим серию исследований, проведенных Уолтером Мишелем — ученым, занимающимся психологией личности, — и его коллегами. Мишель интересовался вопросом, почему дети бывают нетерпеливы, почему они иногда хотят чего-то «прямо сейчас» и почему им трудно ждать. Мишель придумал конструкт, который назвал «задержка удовольствия» и попытался разработать для него подходящий способ измерения. Его исследовательская программа показала, что задержка удовольствия валидна как конструкт и соответствует общей когнитивно-социальной теории личности, а также валидность разработанных им измерений.

Одно простое измерение, придуманное Мишелем, состояло в том, чтобы попросить детей выбрать между маленькой наградой, доступной в этот же момент, и большей, но при условии, что ее выдадут спустя некоторое время. Мишель предположил, что если неспособность к задержке удовольствия является неотъемлемой особенностью маленьких детей, то старшие дети должны охотнее ждать большую награду. Эти рассуждения привели к очевидному эксперименту (процитировано в Mischel, 1981), в ходе которого дети выполнили небольшое задание, а затем им сказали:

> Я бы хотел раздать каждому из вас по конфете, но сегодня я взял с собой очень мало вот таких (показывая большую конфету). Поэтому вы можете либо взять такую (показывая маленькую конфету) прямо сейчас, или, если хотите, можете подождать такую (показывая), которую я принесу в следующий четверг (неделей позже).
>
> Mischel, 1981, p. 164-165

Результаты подтвердили предположение Мишеля: немедленное (но маленькое) вознаграждение было выбрано 81% 7-летних, 48% 8-летних и 20% 9-летних детей.

Естественно, одного этого эксперимента недостаточно для того, чтобы установить валидность задержки удовольствия как конструкта или процедуры измерений, предложенной Мишелем, поэтому он продолжил разработку серии экспериментов, исследуя возможные взаимосвязи между изучаемым конструктом и другими признанными конструктами. Например, он обнаружил, что дети, решившие

отложить удовольствие, были также эмоционально более зрелыми, сильнее нацелены на успех, вероятность делинквентности для них была ниже, а вероятность стать социально ответственными — выше (Mischel, 1981). Таким образом, на основании ряда исследований, результаты которых подтвердили сделанные предположения, задержка удовольствия была признана валидным конструктом.

Надежность и валидность

Чтобы измерение имело ценность для психологического исследования, оно должно обладать одновременно и надежностью, и валидностью. Надежность важна для формирования уверенности в том, что проводимое измерение приближает к истинному значению, а валидность важна, так как она указывает, что измеряется именно то, что вы предполагаете. Обратите внимание, что валидность предполагает надежность, а обратное не верно. Надежные измерения могут не быть валидными, но валидные должны быть надежными.

Проиллюстрировать это можно на простом примере. Из главы 1 вы получили некоторые сведения о френологии — популярной в XIX в. теории, утверждающей, что можно определить «способности» человека по форме его черепа. После обсуждения проблемы надежности вы можете сделать вывод, что френологические измерения черепа действительно обладали высокой надежностью — расстояние между точкой, расположенной пятью сантиметрами выше вашего левого уха, и другой, находящейся пятью сантиметрами выше вашего правого уха, не сильно изменится от измерения к измерению. Однако утверждение, что при измерении определяется «деструктивная» способность человека, — это совсем другое дело. Мы уверены, что измерение черепа не является валидным измерением деструктивности, поскольку оно не кажется нам разумным (очевидная валидность), на основании его невозможно предсказать проявление агрессивности (критериальная валидность), а также потому, что оно не согласуется с другими исследованиями конструктов, связанных с деструктивностью, таких как импульсивность, или исследованиями функций мозга (конструктная валидность).

Вопрос надежности и валидности имеет также этическую сторону, особенно если результаты измерений используются для принятия решений, касающихся жизни людей. Студентов принимают или нет в учебные заведения, людей берут или нет на работу, кому-то ставят психиатрический диагноз и предоставляют лечение — и все это на основании измерений способностей или оценки особенностей поведения. Если вы ищете работу и ваши оценки по некоторому тесту являются решающим показателем, то вы испытаете законное разочарование, узнав, что тест не был достаточно надежным и валидным.

И последнее замечание. В этой главе понятие валидности обсуждалось в контексте темы измерений. Как вы узнаете из следующей главы, валидными могут быть не только измерения психологических конструктов. Валидность также в более общем смысле характеризует весь исследовательский проект. По отношению к измерениям валидность означает, что используемый инструмент измеряет именно то, что должен. В более широком смысле валидность характеризует правильность проведения экспериментов и проверки гипотезы.

Виды шкал измерений

Результаты измерении характеристик поведения представляются в виде набора чисел. Мы говорим, что кто-то среагировал через 3,5 секунды, получил 120 баллов за /Q-тест или нашел выход из лабиринта третьим по счету. Мы также относим X человек к определенной категории в зависимости от их поведения или обладания некоторыми особенностями. Данные примеры иллюстрируют четыре числовых способа описания явлений, а следовательно, четыре **шкалы измерений.** Ясное понимания этих способов необходимо для дальнейшего обсуждения статистических методов, так как от вида используемой шкалы зависит вариант соответствующего статистического анализа данных. Именно неразбериха со шкалами измерений оказалась столь затруднительна для Дилберта, история которого изображена на рис. 4.4.



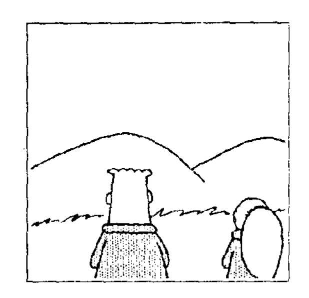




Рис. 4.4. Проблема со шкалой измерений

Номинальная шкала

Иногда численные характеристики, которые мы приписываем явлениям, служат только для того, чтобы разделить эти явления по группам. В этом случае мы используем так называемую **номинальную шкалу** измерений и числа являются не более чем ярлыками категорий (например, 15 мужчин, 25 женщин). В исследованиях с использованием этой шкалы люди обычно разделяются по категориям и далее иодсчитывается количество людей, попадающих в каждую из них. Номинальную шкалу используют, если поставлены, например, такие эмпирические вопросы:

- Среди людей, занимающихся бегом, кто чаще бегает по утрам, а кто по вечерам, мужчины или женщины?
- Если разделить психически неуравновешенных людей на замкнутых и общительных, будут ли интроверты чаще испытывать состояние тревоги, чем экстраверты?
- Если разделить людей на ожидающих сильный электрический шок и на не ожидающих, повлияет ли это их ожидание на тенденцию к аффилиации?

Последний эмпирический вопрос был поставлен в хорошо известном эксперименте по социальной психологии, проведенном Стэнли Шехтером в 1959 г. Половине женщин, участвующих в исследовании, он сказал что в эксперименте будет изучаться психологическая, реакция на сильный электрический шок, а остальным — что разряд будет очень слабым и станет щекотно. После того как женщин в обеих

группах проинформировали о предстоящем испытании, их попросили подождать 10 минут, пока настраивается оборудование, и спросили, как они предпочитают ждать, вместе с другими, в одиночестве, или же не имеют в этом отношении предпочтений. В действительности Шехтера интересовал ответ именно на этот вопрос, и эксперимент был окончен после того, как участницы высказались.

Шехтер изучал влияние состояния тревожности на конструкт, названный аффилиацией (желание находиться вместе с другими). Каковы же были результаты? Те, кто ожидал сильного шока, предпочитали ждать вместе с другими, а те, кто ожидал слабого — в основном не имели предпочтений. Этот эксперимент иллюстрирует применение номинальной шкалы, поскольку Шехтер разделил все данные на шесть категорий и подсчитал количество людей, попадающих в каждую из них в зависимости от того: а) какой силы электрический разряд они ожидали получить и б) как они ответили на ключевой вопрос. Количество людей («подсчет частости») в каждой категории показано в табл. 4.2. Информация также переведена в проценты, что является обычным при использовании номинальных данных.

Таблица 4,2

Использование номинальной шкалы: исследование Шехтером аффилиации

	Количество людей, выразивших следующие предпочтения		
	Ждать в одиночестве	Ждать вместе с другими	без предпочтений
Ожидающие сильный шок	3 (9%)	20 (63%)	9 (32%)
Ожидающие слабый шок	2(7%)	10(33%)	18(60%)

Порядковая шкала

Порядковая шкала измерений представляет собой набор порядковых мест. Например, в списках успеваемости в колледжах обычно отражаются места студентов в группе: 1е, 2е, 3е, 50е и т. д. Из такого списка можно узнать, что один студент получает более высокие оценки, чем другой. Однако относительная позиция — это все, что можно узнать. Эд, Тед и Фред будут занимать соответственно 1-е, 2-е и 3-е места в каждом из двух приведенных ниже случаев, несмотря на то что оценки Эда заметно выше оценок Фреда и Теда только во втором случае:

Случай 1	Случай 2
Средний балл Эда = 4,0	Средний балл Эда = 4,0
Средний балл Фреда = 3,9	Средний балл Фреда = 3,6
Средний балл Теда = 3,8	Средний балл Теда = 3,5

Исследования, в которых используется порядковая шкала, пригодны для разрешения таких эмпирических вопросов:

- Если ребенок расположит пять игрушек в ряд в порядке убывания привлекательности, а затем ему дадут игрушку, занимающую третье место в ряду,

то увеличится или уменьшится порядковый номер этой игрушки в ряду спустя неделю игры с нею?

- Изменяется ли ранжирование студентами авторов учебников по естественнонаучным и гуманитарным дисциплинам, если они знают пол авторов?
- Как молодые и пожилые люди оценят и расположат в ряд 10 кинофильмов с разным количеством сексуальных сцен и сцен насилия?

Хороший пример использования порядковой шкалы — исследование Корна, Девиса и Девиса (Korn, Davis, and Davis), проведенное в 1991 г. Историков психологии и заведующих кафедр попросили назвать 10 психологов, чьи работы, по их мнению, наиболее важны, и расположить их в ряд. Требовалось составить два ряда: один из десяти психологов, лучших «во все времена», и другой — из современников. Затем результаты обобщили. Кто же был первым в списке? И историками, и заведующими кафедр самым выдающимся современным психологом был признан Б. Ф. Скиннер. При этом заведующие кафедр назвали Скиннера и лучшим психологом в истории, а историки, склонные предпочитать психологов более ранних периодов, поместили Скиннера на девятое место, а на первое — Вундта.

Интервальная шкала

В большинстве психологических исследований используется либо интервальная шкала измерений, либо шкала отношений. Интервальная шкала дополняет идею ранжирования принципом равных интервалов между ранжируемыми явлениями. Наиболее распространенный пример использования интервальной шкалы — психологические тесты личности, установок и способностей. Например, результаты теста интеллекта обычно представляются подобным образом. Некто, имеющий *IQ* 120, предполагается более умным (предположим, что *IQ* определяет умственные способности), чем тот, чей *IQ* равен ПО. Более того — и в этом заключается отличительная особенность интервальной шкалы — разница в интеллектуальных способностях двух людей с *IQ* 120 и 110 предполагается такой же, как разница между людьми с *IQ* 110 и 100. Другими словами, предполагается, что каждый шаг увеличения значения *IQ* отражает одинаковую степень увеличения интеллекта (интервалы равные). Обратите внимание на слово «предполагается» — некоторые психологи считают, что *IQ* (а также результаты большинства тестов личности) строятся по принципу шкалы отношений; они утверждают, что довольно трудно, если вообще возможно, в этом случае предполагать равные интервалы. Но большинство ученых считают, что *IQ* относится к интервальной шкале, хотя говорят так отчасти по прагматическим причинам: психологи любят использовать интервальную шкалу и шкалу отношений в основном потому, что структура данных этих шкал позволяет проводить более сложный статистический анализ.

В описанном ранее исследовании булимии использовались различные виды измерений (например, по шкале восприятия социальной поддержки), являющихся примером интервальной шкалы. Во вставке 4.2 рассматривается классическая серия экспериментов, в которых интервальная шкала использовалась при попытке показать влияние телосложения на тип личности.

ВСТАВКА 4.2

Классические исследования — определение соматотипа: когда 7 - 1 - 1 встречает 1 - 1 - 7

Вы уже знаете, что френологи считали, что существует взаимосвязь между определенными физическими характеристиками (форма черепа) и свойствами личности. Френология кажется нам сегодня нелепой, но идея наличия взаимосвязи между физическими особенностями и личностью имеет долгую историю. Более систематическую попытку изучения такой взаимосвязи предпринял Вильям Шелдон в 1940 и 1942 гг.

В работе Шелдона впечатляет его попытка определить телосложение при помощи точной шкалы измерений. Изучив 4000 фотографий обнаженных мужчин (студентов колледжа), он и его исследовательская группа разработали классификацию типов телосложения с помощью трех интервальных шкал, имеющих по 7 делений. Каждая шкала отражала степень выраженности у мужчины одного из идеальных антропометрических типов: эндоморфного (тучный), мезоморфного (мускулистый) и эктоморфного (худой). Предполагалось, что каждый человек несет в себе все три типа, выраженные в разной степени, обычно с преобладанием одного из них. Так, очень полный человек обозначался как 7-1 - 1 , очень худой - как 1 -1 - 7 , а Арнольд Шварценеггер - как 1 - 7 - 1 . 4 - 4 - 4 обозначало гармоничное телосложение. Набор чисел, соответствующих конкретному человеку, назывался его «соматотипом».

После измерений соматотипа Шелдон занялся определением типов личности, которые также, по его мнению, можно разделить на 3 категории и оценить по 7-балльной интервальной шкале, обобщающей результаты различных тестов личности. Он обозначил категории как «висцеротоническую», «соматотоническую» и «церебротоническую». Висцеротоники были в основном общительными, уравновешенными, двигались медленно, любили развлекаться и очень любили поесть, соматотоники были агрессивными, эгоистичными и любили риск, а церебротоники были замкнутыми, застенчивыми, предпочитали одиночество и любили решать умственные задачи.

Заключительным шагом Шелдона была проверка взаимосвязи соматотипа и типа личности. Вас не удивит, что чаще всего вместе появлялись следующие пары:

эндоморф-висцеротоник;
мезоморф- соматотоник;
эктоморф- церебротоник.

Шелдон считал, что тип телосложения является причиной формирования у человека определенного типа личности, но критики его теории указали, что взаимосвязь здесь не такая прямая и может проявляться по-разному. Эндоморфное телосложение может явиться причиной любви к еде, но разве не может сильное пристрастие к еде сформировать эндоморфа?

Вокруг работы Шелдона велись бурные споры, его теория была в целом опровергнута, но для нас эта теория важна тем, что она - классический пример попытки количественно определить человеческий характер и связать его с внешним видом человека, определяемым по шкале телосложений, в данном случае по интервальной шкале. Возможно, вы обратили внимание на то, что в описании работы Шелдона речь идет только о мужчинах. Это потому, что Шелдон только их и изучал. Но почему это так? Вспомните, что его процедура включала изучение фотографий 4000 обнаженных мужчин. По-видимому, он хотел повторить исследование и изучить также и женщин, но засомневался из-за «санкций против фотографирования обнаженных женщин (студенток колледжа)» (Bavelas, 1978). Далее в этой главе вы узнаете об «удобной» выборке и о том, что одни выборки могут быть более пригодны, чем другие.

Необходимо отметить, что на интервальной шкале отметка ноль — это рядовая отметка, не означающая отсутствия измеряемого признака. Стандартным примером является температура. Ноль градусов не означает отсутствия температуры, а лишь является делением шкалы, которое означает «наденьте свитер». Аналогично в тесте на застенчивость, оценки которого варьируются от 0 до 20, значение 0 — это просто нижнее деление на шкале и не означает полного отсутствия застенчивости.

Шкала отношений

В случае **шкалы отношений** принципы упорядоченности и равных интервалов перенесены из структур порядковой и интервальных шкал, но в отличие от них шкала отношений имеет точку *истинного нуля.* На шкале отношений отметка ноль означает полное отсутствие измеряемого признака. Например, если количество ошибок для крысы, пробежавшей лабиринт, имеет значение ноль, то это означает отсутствие неправильных поворотов. Шкала отношений обычно используется при исследованиях, в которых проводятся измерения физических величин, таких как расстояние, вес или время. Примеры, посвященные процедуре привыкания и оценке скорости реакции, приведенные ранее в этой главе, иллюстрируют использование шкалы отношений.

«Кого» измерять: выборки

Кроме решения о том, что измерять при проведении психологических исследований, необходимо также решить, кого просить участвовать в исследовании и чье поведение будет оцениваться. Для этого есть два подхода: вероятностная и простая (невероятностная) выборки.

Вероятностная выборка

Эта стратегия используется, если поставлена задача узнать что-либо конкретное об определенной группе людей. Группа в целом называется популяцией, а любая из ее подгрупп называется **выборкой.** Иногда бывает возможно изучать всех членов популяции. Например, если вы хотите изучить отношение всех студентов вашей группы по экспериментальной психологии к экспериментам с участием животных и не хотите делать выводы для находящихся за пределами группы, вы можете опросить всех ее членов. В этом случае размер популяции будет равен размеру вашей группы. Однако, как вы можете догадаться, интересующие исследователей популяции обычно слишком велики для того, чтобы изучать каждого их члена. Поэтому из этой популяции необходимо отбирать подмножество, или выборку.

Хотя вся популяция может и не рассматриваться в исследовании, ученым требуется делать выводы именно о популяции, а не только о выборке. Поэтому важно, чтобы выборка отражала особенности популяции в целом. Если это так, выборка считается **репрезентативной,** если нет — **нерепрезентативной.** Если вы хотите изучить восприятие студентами учебы в колледже, то будет серьезной ошибкой выбрать для исследования только тех, кто живет в студенческом общежитии. Из-за того что отношение прочих студентов может отличаться от отношения живущих в общежитии, результаты вашего опроса могут оказаться необъективными и сместиться в пользу последних (более подробно о процедуре опроса вы узнаете в главе 12).

Одним из наиболее известных исторических примеров нерепрезентативной выборки является случай, произошедший во время президентских выборов в 1936 г. Успешно прогнозируя события нескольких предыдущих выборов, журнал «Литрери Дайджест» попытался предсказать результаты выборов, послав 10 миллионов пробных бюллетеней своим подписчикам, людям, выбранным по телефонным книгам всей страны, и людям из регистрационных списков автомобилей (Sprinthall, 2000). Около 25% бюллетеней (почти 2,5 миллиона) вернулись в журнал заполненными. Из людей, заполнивших бюллетень, 57% выбрали кандидата от республиканской партии Альфа Лэндона, а 40% выбрали действующего президента Франклина Рузвельта. На действительных выборах победил Рузвельт, набрав более 60% голосов. Как вы думаете, почему выборка была нерепрезентативной?

Хотя редакторы журнала знали, что их подписчики в основном принадлежат к верхней прослойке среднего класса и одновременно являются республиканцами, они подумали, что, расширив выборку за счет людей, выбранных из телефонных книг и регистрационных списков, они сделают ее более репрезентативной. В действительности они просто набрали еще больше республиканцев. В разгар Великой депрессии позволить себе иметь телефоны и автомобили могли лишь принадлежащие к верхней прослойке среднего класса и верхнему классу люди, а они в большинстве своем были республиканцами, а не демократами. Поэтому на самом деле в своем опросе журнал спрашивал республиканцев о том, кому они отдадут свои голоса.

Возможно, вы заметили еще один недостаток опроса в «Литрери Дайджест». Большое количество бюллетеней вернулось, и составители журнала были вполне уверены в правильности предсказания о победе Лэндона, так как полученные данные отражали взгляды значительного числа людей — около 2,5 миллионов. Но заметьте, что помимо того, что полученные бюллетени составили только одну четверть от посланных, все эти бюллетени были отправлены теми людьми, которые *сами решили* это сделать. Так что люди, приславшие бюллетени обратно в журнал, были не просто республиканцами, но республиканцами, пожелавшими сделать свои взгляды общеизвестными (не кажется ли вам, что в свете этого 57% выбравших Лэндона — это довольно мало?).

Проблема самовыбора нередко возникает при опросах, проводимых популярными журналами, а также в случае вопросов, задаваемых своим читателям такими людьми, как Энн Лэндер, чтобы узнать их мнение по поводу отдельных тем. Проводится опрос, затем спустя месяц или два результаты, полученные от тех, кто прислал обратно заполненный опросник, представляются в виде отчета, подразумевающего их валидность. При этом человек, делающий отчет об исследовании, старается произвести впечатление числом ответов, а не репрезентативностью выборки. Пример такого явления — отчет о женской сексуальности (Hite, 1987), в котором кроме прочего утверждалось, что более 90% женщин, состоящих в браке, чувствуют себя эмоционально ущемленными. В ответ на замечание, что опрос был послан лишь отдельным женским организациям и что заполненные опросники прислали обратно только 4,5% из 100 000 женщин, автор просто заявила, что 4500 человек для нее достаточно (как 2,5 миллиона человек было достаточно для «Литрери Дайджест»).

Обладая научным мышлением, вы должны весьма скептично относиться к утверждениям, сделанным на основании такого рода выборок. Если вы хотите еде-

лать правильный вывод о конкретной популяции, необходимо использовать выборку, отражающую популяцию в целом, и самостоятельно осуществлять отбор, не полагаясь на тех, кто решил вам ответить самостоятельно.

Случайная выборка

Основной вид вероятностной выборки — **простая случайная выборка.** По сути, она означает, что каждый член популяции с равной вероятностью может попасть в выборку. Например, чтобы сделать случайную выборку из 100 студентов вашего учебного заведения, вы можете сложить бумажки с именами всех студентов в большую шляпу, а затем достать из нее 100 бумажек. В действительности процедура немного более сложнее и обычно проводится с использованием таблицы случайных чисел, образец которой представлен в приложении *D* Чтобы узнать, в чем заключается эта процедура, изучите пример, приведенный в табл. 4.3, — в нем описывается, как использовать случайные числа для выбора 5 человек из популяции в 20 человек.

Простая случайная выборка — эффективный практический способ построения репрезентативной выборки. Иногда его также используют по этическим причинам. В случае если только небольшая группа людей может получить выгоду или понести издержки и нет других разумных оснований для принятия решения, лучше всего использовать метод случайной выборки. Известный пример такой выборки — лотерея для отбора людей, призывавшихся на военную службу, проводившаяся в 1969 г. в разгар Вьетнамской войны. Чтобы все было честно, 365 дней года должны были иметь равную вероятность стать первым, вторым, третьим и т. д. К сожалению, действительная процедура имела некоторые недостатки (Kolata, 1986). В большой барабан поместили капсулы, соответствующие каждому из дней года, при этом одновременно опускали все капсулы одного месяца. Первыми опустили январские капсулы, затем февральские и т. д. Барабан прокрутили, чтобы перемешать капсулы, но, по-видимому, не очень тщательно, потому что когда стали выбирать дни, капсулы, опущенные последними, в большинстве своем выбирались первыми. В то время лучше было не отмечать свой день рождения в декабре.

Простая случайная выборка имеет два недостатка. Во-первых, вам может понадобиться отразить в выборке определенные особенности популяции, а во-вторых, процедуру может быть невозможно применить на практике в случае, если популяция очень велика. Сможете ли вы получить список всех людей, проживающих в Соединенных Штатах, если вам потребуется сделать простую случайную выборку североамериканцев? Первую проблему решает расслоенная выборка, а вторую — гнездовая.

Таблица 4.3

Создание случайной выборки с помощью таблицы случайных чисел

Задача: Сделайте случайную выборку в 5 человек из популяции в 20 человек

Шаг 1 Обозначьте номерами от 01 до 20 всех членов, составляющих популяцию

Шаг 2 Перейдите к таблице случайных чисел. Ниже приведена часть такой таблицы:

2	2	1	7	6	8	6	5	8	4	6	8	9	5
1	9	3	6	1	7	5	9	4	6	1	3	7	9
1	6	7	7	2	3	0	2	7	7	0	9	6	1
7	8	0	3	7	6	7	1	6	1	2	0	4	4
0	3	2	8	1	2	2	6	0	8	7	3	3	7

Окончание табл. 4.3

Шаг 3	Выберите число из таблицы, с которого вы начнете поиск. Оно может быть любым, единственное, чего нельзя делать, — это начинать два отдельных поиска с одного и того же числа. Предположим, вы начинаете с верхнего числа в третьем столбце слева. Это число 1. Вам потребуется выбирать пары чисел, образующие двузначные числа, так как популяция состоит из людей, пронумерованных от 01 до 20. Поэтому в качестве стартовой точки можно использовать число 1 и стоящее за ним число 7
Шаг 4	Поиск выборки начинается с числа 17, попадающего в промежуток от 01 до 20. Таким образом, первый человек в вашей выборке имеет номер 17
Шаг 5	Продолжайте поиск вниз по двум столбцам пока не найдете пять человек с номерами от 01 до 20. Ниже опять приведена таблица с разбивкой столбцов для удобства подбора числовых пар. Пять выбранных пар подчеркнуты, а участок таблицы, в котором нужно производить поиск, выделен жирным шрифтом. Стрелки указывают направление поиска:

2	2	**1**	**7**	**6**	**8**	**6**	**5**	**8**	4	6	8	9	5
1	9	**3**	**6**	**1**	**7**	**5**	**9**	**4**	6	1	3	7	9
1	6	**7**	**7**	**2**	**3**	**0**	**2**	**7**	7	0	9	6	1
7	8	**0**	**3**	**7**	**6**	**7**	**1**	**6**	1	2	0	4	4
0	3	**2**	**8**	**1**	**2**	**2**	**6**	**0**	8	7	3	3	7

Таким образом, выборка состоит из членов популяции с номерами 17,3,12, 2 и 8. Обратите внимание, что номера, превышающие 20 (например, 36), пропускаются и что если номер повторяется дважды, его повторно не выбирают

Расслоенная выборка

Предположим, вы хотите узнать отношение к абортам в вашем кампусе, а популяция вашего учебного заведения состоит из 5000 человек, из которых 4000 — женщины. Для этого вы решаете отобрать 100 студентов. Если вы используете простую случайную выборку, то, вероятно, в вашей выборке будет больше женщин, чем мужчин, но их соотношение не будет точно таким же, как в популяции. Вашей задачей является создание выборки, действительно отражающей состав популяции, а для такого вопроса, как аборт могут быть важны различия во мнениях между мужчинами и женщинами. Поэтому если в выборки окажется больше мужчин, то она может неверно отобразить отношение к этому вопросу в вашем кампусе. В подобном случае неплохо было бы заранее решить, что если 80% популяции — женщины, то они должны составлять ровно 80% выборки.

В расслоенной выборке точно отображается *соотношение* важных подгрупп популяции. В приведенном выше примере из списка женщин случайным образом должны были быть отобраны 80 человек, **а** из списка мужчин — 20.

Обратите внимание, что в таком случае исследователю требуется решить, на сколько уровней делить популяцию. В случае опроса по поводу аборта были отобраны в соответствии с их реальным соотношением мужчины и женщины. Должны ли быть пропорционально представлены в выборке также все четыре студенческих потока? А как насчет католиков и протестантов? Право- и леворуких? Оче-

видно, что исследователь должен где-то остановиться. Для расслоения выборки одни характеристики (религия) являются более важными, чем другие (право- или леворукость). Исследователь должен разумно подойти к решению этого вопроса в соответствии с условиями исследования или целями экспериментов.

Кластерная выборка

Метод расслоенной выборки весьма эффективен, но он не решает проблему построения выборки из большой популяции, для которой обычно невозможно получить полный список членов. Эту проблему решает **кластерная выборка,** которую часто используют национальные службы опроса общественного мнения. С помощью этого подхода исследователи случайным образом выбирают группу людей (кластер), имеющих определенную особенность. Используя этот метод, можно проводить опрос в кампусе большого университета. Если исследователь хочет получить репрезентативную выборку студентов, а применить расслоенную выборку невозможно, то альтернативой будет список базовых предметов данного учебного заведения. Студенты разных специальностей, посещающие занятия по одному предмету, образуют группу. Если в вузе преподают 40 базовых предметов, исследователь может выбрать 10 из них, а затем провести опрос всех студентов, посещающих эти курсы.

Если выбранные группы слишком велики, исследователь может выбрать из них более мелкие подгруппы. Предположим, вы хотите узнать, нравится ли студентам жить в многоэтажных общежитиях вашего кампуса, которые вы операционально определили как любые общежития, имеющие восемь этажей и более. Допустим, в вашем кампусе есть 15 таких зданий, в которых проживают в общей сложности **6000** студентов. Воспользовавшись методом гнездовой выборки, сначала можно выбрать шесть зданий (каждое здание = группа), а затем для каждого здания случайным образом выбрать три этажа и включить в выборку всех проживающих на выбранных этажах в выбранных общежитиях (предположим, около 40 человек на этаже). Таким образом, общее число выборки будет 720 человек (40 х 3 х 6). Обратите внимание, что в данном случае можно также использовать некоторые элементы расслоенной выборки. Если из 15 общежитий 10 женских и 5 мужских, то можно составить группу, отражающую это соотношение: 4 женских общежития и 2 мужских.

Простая выборка

Прочитав приведенную выше информацию, вы можете решить, что без вероятностной выборки качество исследования снизится. Это действительно так, если целью проекта является точное описание конкретных особенностей некоторой популяции на основании изучения ее части. Но большинство исследований по экспериментальной психологии не таковы и обычно цель исследования заключается в изучении отношений между переменными, например: использование изображений улучшает запоминание; наблюдение агрессии приводит к агрессивному поведению; с увеличением количества свидетелей происшествия снижается число желающих оказать помощь и т. д. Конечно, можно надеяться на вероятность того, что результаты таких исследований будут пригодны не только для людей, участвующих в них, но исследователи просто предполагают, что если выявлена сильная взаимосвязь, то она обязательно проявится у большинства людей, принадлежащих к определенной популяции, вне зависимости от способа их отбора. Будет ли это предположение

истинным, естественно, зависит от стандартных процессов повторения и дополнения эксперимента, рассмотренных в главе 3. Поэтому, например, при изучении границ кратковременной памяти у взрослых нет необходимости делать случайную выборку — подойдет практически любая группа взрослых людей.

Удобная выборка

Это чаще всего используемый (и действительно наиболее удобный) вид простой выборки. Чтобы создать **удобную выборку,** исследователь просто использует добровольцев из людей, соответствующих основным требованиям исследования. Обычно это первокурсники или второкурсники из «фонда испытуемых» — студенты, занимающиеся общей психологией, которых попросили поучаствовать в одном или нескольких исследованиях (возможно, вам тоже приходилось участвовать в такой работе). Более подробно об этических аспектах «фонда испытуемых» см. вставку 5.2.

Иногда при этом требуется участие определенного типа людей (иногда такая выборка называется «целевой»). Например, когда Стэнли Милгрэм в первый раз набирал участников для исследования явления подчинения, он дал в местную газету объявление о наборе добровольцев. Он намеренно не стал использовать студентов, так как считал, что они могут оказаться «слишком однородной группой... (он) хотел исследовать разных людей, принадлежащих к различным слоям общества» (Milgram, 1974, p. 14). Он мог бы попробовать более сложный подход и сделать расслоенную или кластерную выборку, но не было гарантии, что выбранные им люди согласятся прийти в лабораторию. Кроме факта, что не для всех исследований требуется делать выборки, точно отражающие популяцию, способ, выбранный Милграмом, показывает, что часто имеются прагматические основания для использования простой выборки.

Статистический анализ

Первое положение хорошо известного каждому гражданину Соединенных Штатов руководства «Неисхоженная дорога» гласит: «Жизнь трудна» *(The Road Less Traveled,* Peck, 1978, p. 15). Верить в это начинают многие студенты, приступающие к изучению статистики (они легко узнают себя на рис. 4.5). Я не стану убеждать вас в том, что занятия статистикой подобны отдыху на пляже во Флориде, но, надеюсь, вы увидите, что часть удовольствия от психологического исследования приносят анализ информации, которую вы так усердно собирали, и выяснение, *действительно* ли исследование показало то, что, как вы думаете, оно показало. Я видел, как взрослые и ответственные люди затаив дыхание ожидали результатов статистического анализа, а затем поникали головой или прыгали от восторга, едва взглянув на волшебные числа. Так что если у вас появилась страсть к психологическим исследованиям, тонкости статистического анализа вызовут у вас немало эмоций.

Моя задача состоит в том, чтобы познакомить вас со статистическим образом мысли и теми видами статистического анализа, с которыми вы встретитесь, проводя психологические исследования. Хорошо, если вы уже изучили курс статистики, но если нет, то вам необходимо сделать это как можно быстрее, особенно если вы хотите поступать в аспирантуру на связанную с психологией специальность. Хотите верьте, хотите нет, но в списке называемых аспирантурами дисциплин, ко-

торые обязательно должны изучить поступающие, статистика занимает первое место, а второе — методы исследований (вспомните описанное в главе 1 исследование, проведенное Норкроссом, Ханичем и Терранова в 1996 г.). Я расскажу вам об основных статистических методах, но это не заменит полного курса.



Рис. 4.5. Распространенное представление о статистике

Информацию о том, как выбрать подходящий метод статистического анализа и как проводить некоторые обычные процедуры обработки данных, вы найдете в приложении *С.* По ходу книги время от времени вам будут также встречаться темы, касающиеся статистики, поскольку разработку психологического исследования невозможно отделить от анализа этого исследования. В этой главе вы узнаете о различиях между описательной статистикой и статистикой вывода, а также о логике проверки гипотез.

Описательная статистика и статистика вывода

Два самых общих вида статистической обработки данных — это описательная статистика и статистика вывода. Различия между ними соответствуют различию между выборкой и популяцией. Говоря просто, **описательная статистика** обобщает данные, собранные на выборке участников занятых в вашем исследовании, а **статистика вывода** позволяет вам делать такие выводы об этих данных, которые могут быть применены к популяции в целом.

Описательная статистика

По сути, методы описательной статистики позволяют вам свести огромное количество чисел, смысл которых невозможно охватить сразу, к очень небольшому набору, значение которого понять гораздо легче. Описательная статистика включает оценку общей тенденции, изменчивости и взаимосвязей, представленных как численно, так и наглядно (в виде графиков). В этой главе мы рассмотрим основные процедуры оценки общей тенденции и изменчивости. Оценка взаимосвязей (вычисление коэффициентов корреляции) будет описана в главе 9.

Для иллюстрации оценки общей тенденции и изменчивости рассмотрим данные гипотетического исследования памяти, в котором 20 человек запоминали, а затем пытались воспроизвести список из 25 слов. Каждое представленное ниже число соответствует количеству слов, запомненных каждым из 20 участников:

16	18	19	19
17	19	15	21
14	16	15	17
17	20	17	15
18	17	18	18

Сразу видно, что обобщение результатов этого исследования требует чего-то большего, чем простой демонстрации набора из 20 чисел. Например, можно попытаться вычислить *типичную* оценку, или так называемую «общую тенденцию». Чаще всего психологи-исследователи определяют общую тенденцию вычисляя среднее арифметическое. Для этого складывают все оценки и делят полученную сумму на общее количество оценок:

$$\bar{X} = \frac{\sum x}{n};$$

где X = среднее арифметическое; ΣX = сумма отдельных оценок; n = количество оценок в примере.

В случае данных, собранных при исследовании памяти, получаем:

$$\bar{X} = \frac{\sum x}{n} = (16 + 17 + 14 + ... + 18)/20 = 346/20 = 17{,}30.$$

Два других способа нахождения общей тенденции — это вычисление медианы и моды. **Медиана** представляет собой оценку, находящуюся строго в середине набора оценок. Одна половина оценок выше, а другая — ниже значения медианы. Для определения медианы в первую очередь нужно составить последовательность оценок, от наименьших к наибольшим. В случае данных, собранных при исследовании памяти, последовательность будет следующая:

14 15 15 15 16 16 17 17 17 17 17 18 18 18 18 19 19 19 20 21
⇑

Далее нужно определить **местоположение медианы** — позицию в последовательности оценок, где проходит медиана (Howell, 1997). Это вычисляется по формуле:

$$\text{местоположение медианы} = \frac{n+1}{2}.$$

Для данных из исследования памяти местоположение медианы следующее: (20+1)/2 = 10,5. Это означает, что она лежит посередине (0,5) между 10-ми 11-м номерами в последовательности. Считая слева направо, видим, что и 10-й, и 11-й номера — это число 17 (я отметил это место в показанной выше последовательности знаком 1Ї). Медиана является точной серединой набора оценок: с каждой стороны от нее лежит по 10 чисел.

Иногда медиану используют, если набор оценок содержит одну или две, сильно отличающихся от остальных. В такой ситуации среднее арифметическое дает искаженное представление о типичной оценке. Предположим, к примеру, что пять преподавателей с вашего факультета психологии получили следующие оценки *IQ*. 93,81,81,95 и 200 (последняя оценка вероятно принадлежит преподавателю методов исследований). Среднее арифметическое оценок *IQ,* равное 110 (вы можете проверить), дает ложное представление о том, что в целом преподаватели психологического факультета имеют умственные способности заметно выше среднего. Медиана в данном случае позволяет лучше оценить типичную /Q-оценку. Местоположение медианы равно (5+1)/2 = 3, а в последовательности оценок третье число равно 93:

81 81 93 95 200

1\

Очевидно, что медиана оценок *IQ,* равная 93, гораздо лучше отражает обычный уровень интеллектуальных способностей на данном гипотетическом факультете психологии.

Мода — это значение, чаще всего встречающееся в наборе оценок. В приведенном выше примере значение моды равно 81. Мода гипотетических оценок теста памяти равна медиане: число 17 встречается 5 раз, т. е. чаще всех других чисел. Так как в данных теста памяти отсутствуют необычно высокие или низкие оценки, значения среднего арифметического (17,3), медианы (17) и моды (17) довольно близки друг другу, и каждое из них дает верное представление об общей тенденции.

Очевидно, что оценка общей тенденции требует суммирования данных. Менее очевидна, но не менее важна необходимость анализа изменчивости набора оценок. Предположим, вы — гольфер-профессионал и собираетесь вести занятия в местном клубе для двух групп: в 8:00 и 9:00. Вы измерили их способности, определив среднюю оценку для 9 лунок. Ниже приведены полученные вами данные:

Группа, занимающаяся в 8:00:	50	52	58	46	54
Группа, занимающаяся в 9:00:	36	62	50	72	40

Обратите внимание, что среднее арифметическое для каждого набора оценок гольферов равняется 260/5 = 52 ударам. Профессионалу будет о чем поговорить с каждым членом обеих групп. В группе, занимающейся в 8:00, оценки близки друг к другу и все ее участники имеют примерно одинаковый уровень способностей, однако вторая группа не настолько благополучна — оценки в ней варьируются от 36 (довольно хорошо) до 72 (ай-ай-ай!). Понятно, что перед началом занятий гольфер-профессионал предпочел бы знать не только среднюю оценку группы.

Самый простой и весьма приблизительный способ оценить изменчивость — это найти **разброс** — разницу между наибольшей и наименьшей оценками в группе. Диапазон данных для приведенного ранее теста памяти равен 7 (21-14). Разброс оценок 8-часовой группы в примере с занятиями гольфом равен 12 (58 - 46), а разброс оценок 9-часовой — 36 (72 - 36). Разброс дает грубую оценку изменчивости и показывает лишь разницу между крайними значениями. Более сложный способ измерения изменчивости — нахождение стандартного отклонения. Этот способ чаще всего применяется при создании сводного отчета о собранных данных.

Стандартное отклонение для выбранного набора оценок — это среднее значение, на которое оценки данного распределения отклоняются от среднего арифметического этих оценок. В табл. 4.4 показаны два способа вычисления стандартного отклонения. Первый способ вытекает непосредственно из определения и позволя--т лучше понять сущность стандартного отклонения. Второй — это вычислительная формула, использовать которую удобнее при работе с калькулятором. Для при--ера с гипотетическим исследованием памяти стандартное отклонение равняется 1,81 словам. Для примера с занятиями гольфом стандартное отклонение для 8- часовой группы равняется 4,47, а для 9-часовой — 15,03 ударам.

Таблица 4,4

Вычисление стандартного отклонения

Если вы пользуетесь пакетом статистических программ *SPSS* или *SAS,* практически любой анализ будет включать нахождение стандартного отклонения. Кроме того, большинство калькуляторов способны к вычислению основных статистических функций, в том числе стандартного отклонения. Поэтому вам, может быть, не придется подсчитывать на бумаге. Однако у вас может оказаться устаревший калькулятор или увлеченный преподаватель, который будет разделять мои представления о том, что подсчет вручную дает более глубокое понимание происходящего. Есть два способа вычислить стандартное отклонение. Первый — использовать так называемую «формулу отклонения». Внимательно изучив ее, вы лучше поймете суть стандартного отклонения, которое по определению является приблизительной характеристикой средней величины отклонения каждой оценки от среднего арифметического. Ниже показано, как найти стандартное отклонение для 20 оценок, полученных при исследовании памяти.

Шаг 1 Вычислите среднее арифметическое:

$$\bar{X} = \frac{\sum X}{n} = \mathbf{17{,}3}.$$

Шаг 2 Вычислите оценки отклонения, каждую возведите в квадрат и найдите их сумму. Оценки отклонения (*x* малое) находятся вычитанием среднего арифметического из каждой оценки (*X* большое). Таким образом, *x*- *X*- *X*. Возведение в квадрат предотвращает появление отрицательных чисел:

X	$\bar{X}$	x	x^2
16	17,3	−1,3	1,69
17	17,3	−0,3	0,09
14	17,3	−3,3	10,89
...	...	...	...
18	17,3	−0,7	0,49
			$\sum x^2$= **62,20**

Шаг 3 Вычислите стандартное отклонение (СО):

$$CO \approx \sqrt{\frac{\sum x^2}{n-1}} = \sqrt{\frac{62{,}2}{20-1}} = \sqrt{3{,}27} = \mathbf{1{,}81}.$$

По формуле отклонения найти стандартное отклонение довольно просто, но для калькулятора она не совсем подходит. Более простой способ — использовать так называемую формулу для вычислений, которая математически равнозначна формуле отклонения. Она имеет следующий вид]

$$CO=\sqrt{\frac{\sum X^2-\frac{(\sum X)^2}{n}}{n-1}},$$

а вычисления проходят следующим образом:

Шаг 1 Вычислите $\sum X^2$ и $(\sum X)^2$;

$\sum X^2 = 16^2 + 17^2 + 14^2 + ... + 18^2 = 256 + 289 + 196 + ... + 324 =$ **6048**;

$(\sum X)^2 = (16+17+14+...+18)^2 = 346^2 =$ **119716**.

Шаг 2 $\sqrt{\frac{\sum X^2-(\sum X)^2}{n}} = 6048-\frac{119716}{20} = 6048-5985{,}8 \approx$ **62,2**.

Шаг3 Разделите значение, полученное на шаге 2, на $(n-1)$:

$$\frac{62{,}2}{19} = \mathbf{3{,}27}.$$

Шаг 4 Чтобы получить стандартное отклонение, найдите квадратный корень значения, полученного на шаге 3:

$$CO = \sqrt{3{,}27};$$

$$CO = \mathbf{1{,}81}.$$

Одной из характеристик изменчивости является **дисперсия.** Дисперсия представляет собой число, получаемое в ходе вычисления стандартного отклонения, сразу перед нахождением квадратного корня (3,27 для оценок исследования памяти). Это число редко попадает в отчеты, включающие описание данных, так как оно отражает измеряемую величину, возведенную в квадрат (например, «количество запомненных слов в квадрате»). Однако оно находится в центре вероятно самой известной в психологии процедуры статистики вывода — «дисперсионного анализа». О нем рассказывается в главах 7 и 8, а также более подробно в приложении *С.*

Общая тенденция и изменчивость — это универсальные характеристики, используемые при любом описании данных, но исследователи также изучают и весь набор оценок в целом. Простой просмотр данных малоэффективен, но есть и другие способы организации оценок, с помощью которых можно получить значимую картину результатов. Один из способов представления данных — это **гистограмма.** Гистограмма представляет собой график, показывающий, сколько раз встречается каждая оценка в данном наборе, или, при большом количестве оценок, частоту появления оценок в пределах определенного интервала. Чтобы построить гистограмму, необходимо предварительно построить **частотное распределение** — таблицу, в которой указывается, сколько раз встречается каждая оценка. Частотное распределение оценок, полученных при исследовании памяти, имеет следующий вид:

Оценка	Частота	Частота, обозначенная звездочками
14	**1**	*
15	**3**	***
16	**2**	**

17	**5**	*****
18	**4**	****
19	**3**	***
20	**1**	*
21	**1**	*

Построив таблицу частотного распределения, несложно начертить гистограмму. На оси *X* графика отметьте сами оценки, а на оси *Y* — частоту их появления, а затем постройте соответствующие столбцы графика. Результат должен выглядеть, как показано на рис. 4.6. Обратите внимание, что если взять столбец со звездочками из частотного распределения и повернуть его на 90°, результат будет такой же, как на рис. 4.6.

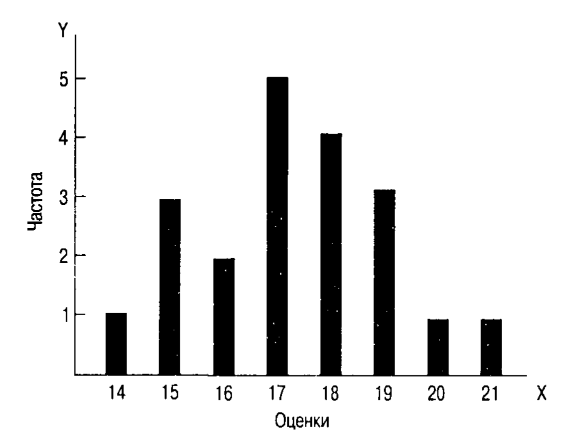


Рис. 4.6. Гистограмма оценок, полученных по тесту памяти

Также следует отметить, что гистограмма выступает вверх в районе середины и уплощается по краям, что приблизительно соответствует распределению оценок для целой популяции, а не только для 20 человек из описанного выше примера. Распределение оценок для популяции представляет собой известную колокообразную кривую, называемую **нормальной кривой,** или нормальным распределением. Вы уже встречались с ней; она представлена на рис. 4.7.

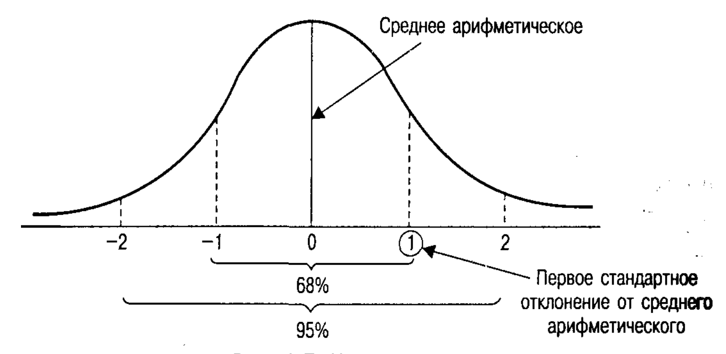


Рис. 4.7. Нормальная кривая

Так же как кривая, построенная для оценок исследования памяти, нормальная кривая представляет собой частотное распределение. Но в отличие от первой она является *нереальным* (или «эмпирическим») распределением оценок конкретной выборки, а *гипотетическим* (или «теоретическим») распределением оценок, которые могут получить члены популяции, если все они примут участие в исследовании. Среднее арифметическое, медиана и мода находятся точно в центре нормального распределения. Важнейшая особенность статистического анализа частотного распределения заключается в том, что если эмпирическое распределение оценок сходно с нормальным распределением, то математические характеристики последнего можно использовать для построения выводов о первом.

Обратите внимание, что на нормальной кривой, показанной на рис. 4.7, я отметил по два стандартных отклонения с обеих сторон от среднего арифметического. Математические характеристики кривой таковы, что около 68% всех оценок для популяции лежат в интервале между двумя первыми стандартными отклонениями, а около 95% — между вторыми. Очевидно, что оценок, попавших за пределы вторых стандартных отклонений, немного — всего 5% от общего количества. Все эти явления можно назвать «статистически значимыми». Запомните данные характеристики распределения, мы к ним очень скоро вернемся.

Кроме частотного распределения и гистограммы есть еще один способ отображения набора данных, который позволяет выявить их особенности. Это **метод стебля и листа** (Turkey, 1977). Чаще всего его используют, когда набор оценок так велик, что частотное распределение или гистограмма были бы очень громоздкими. Например, если вы протестировали 20 испытуемых на застенчивость и полученные ими оценки варьируются от 10 до 70, простое частотное распределение, подобное построенному для данных исследования памяти, будет огромным, а ось *X* гистограммы будет в милю длиной. Проблему можно решить сгруппировав данные по интервалам (10-19,20-29,30-39 и т. д.). Каждый столбец диаграммы будет отралсать количество оценок в пределах определенного интервала. Обратите внимание, что подобная группировка данных приводит к потере некоторой информации. Если шесть человек при тестировании на застенчивость получат оценки между 30 и 39, то все, что вы увидите после такого обобщения, — это один столбец, отображающий частоту, равную шести, и вы не будете знать, какую оценку получил каждый из шести участников. Организовав данные методом стебля и листа, вы сможете получить эту информацию. Метод состоит в следующем. Предположим, что при тестировании на застенчивость 20 человек получены следующие оценки (я выделил жирным шрифтом шесть оценок в пределах от 30 до 39):

49	22	33	46
36	64	39	41
41	68	43	61
36	47	32	49
43	67	37	43

В методе стебля и листа с двухзначными числами «листом» будет наименьший разряд (разряд единиц), а «стеблем» — наибольший (разряд десятков). Таким образом, для первого числа (49) , стеблем будет число 4, а листом число 9. Для числа 36 стебель равен 3, а лист — 6. Для организации стеблей и листов по одноименному методу сначала требуется расположить числа в порядке возрастания, как вы делали при нахождении медианы (числа от 30 до 39 выделены жирным шрифтом). Получаем:

22 32 33 36 36 37 39 41 41 43 43 43 46 47 49 49 61 64 67 68

Далее поместите стебли в левый столбец таблицы, а листы в соответствующие ряды правого столбца, как показано ниже:

Стебли	Листы
2	2
3	236679
4	113336799
5	
6	1478

Повернув таблицу влево на 90° и представив, как заполняются цветом цифры листов, образуя столбцы, вы получите аналог гистограммы для сгруппированных данных. Но обратите внимание, что по сравнению с обычной гистограммой метод стебля и листа обладает заметным преимуществом. На гистограмме, к примеру, в интервале 30-39 будет изображен один столбец, достигающий по шкале *Y* отметки 6. В таблице, построенной методом стебля и листа, вы не только увидите «высоту» оценок в интервале, но также сможете изучать сами оценки. Кроме того, метод стебля и листа позволяет обнаружить оценки, относительно далеко отстоящие от остальных. В приведенном выше примере отсутствие оценок в интервале 50-59 сразу заметно, а четыре оценки в интервале 60-69 выделяются и несколько отстоят от остальных.

В статьях, посвященных результатам исследований, полученных с помощью описательной статистики, встречается три способа представления данных. Во-первых, если необходимо представить лишь несколько чисел (например, значения среднего арифметического и стандартного отклонения для двух экспериментальных групп), можно использовать повествовательное изложение результатов. Во-вторых, значения среднего арифметического и стандартного отклонения можно представить в виде таблицы, а в третьих — наглядно в виде графика. Как строить таблицы и графики, соответствующие стандартам *APA,* вы узнаете из приложения *А,* в котором приведен пример отчета об исследовании. Также, некоторую информацию о построении графиков можно найти в главах 7 и 8. Этический аспект статистического анализа и построения графиков освещается во вставке 4.3.

ВСТАВКА 4.3

Этика — обман с помощью статистики

Все мы были жертвами использования статистической информации для обмана. При этом первыми в ряду жуликов идут политики, а авторы рекламы аспирина занимают близкое к ним второе место. Но и вообще статистикой злоупотребляли достаточно часто для того, чтобы большинство людей стали к ней относиться весьма скептически. Часто можно услышать: «Статистика покажет все, что вы захотите». Действительно ли это так?

Конечно, в каждом случае приходится решать специально, каким образом представлять данные, какие выводы из них сделать. Можно быть вполне уверенным в том, что статьи, которые публикуются в солидных журналах, проверяются и рецензируются специалистами и содержат выводы, основанные на надежных статистических данных. Внимательно следует относиться к статистике, если она используется в более массовых источниках или если ее применяют люди, очевидно стремящиеся убедить вас в истинности собственных убеждений.

Знание о том, как правильно применять статистику, поможет вам распознать сомнительные случаи ее использования. Впервые некоторые моменты, заслуживающие особого внимания, были описаны в прекрасной работе Даррелла Хаффа «Как обмануть с помощью статистики», опубликованной в 1954 г. Работа начинается с цитаты, приписываемой британскому политику Бенджамину Дизраели и выражающей общее мнение по поводу статистики: «Существуют ложь, чудовищная ложь и статистика» (р. 2).

В данной книге мы приведем пример обмана с помощью статистики, в котором был использован способ наглядного представления данных- графики. Предположим, вы читаете об исследовании, в котором сравнивается скорость реакции у мужчин и женщин, и встречаете следующий график (рис. 4.8).



Рис. 4.8. Гистограмма с неразмеченной осью У, отображающая предполагаемое различие скоростей реакции у мужчин и женщин

Такой вид графика Хафф назвал «Ого-графиком» по той очевидной причине, что, когда вы смотрите на него в первый раз, вы восклицаете «Ого! Какая огромная разница!» Однако дело в том, что это различие сильно преувеличено манипуляциями с осью *Y*. Обратите внимание, что она не размечена. Если бы она была размечена, деления выглядели бы, как показано слева на рис. 4.9. Разметка оси V, приведенная на графике справа, гораздо реалистичнее - она дает более верную картину отсутствия значимых различий. Это, скорее всего, подтвердит и статистический анализ.

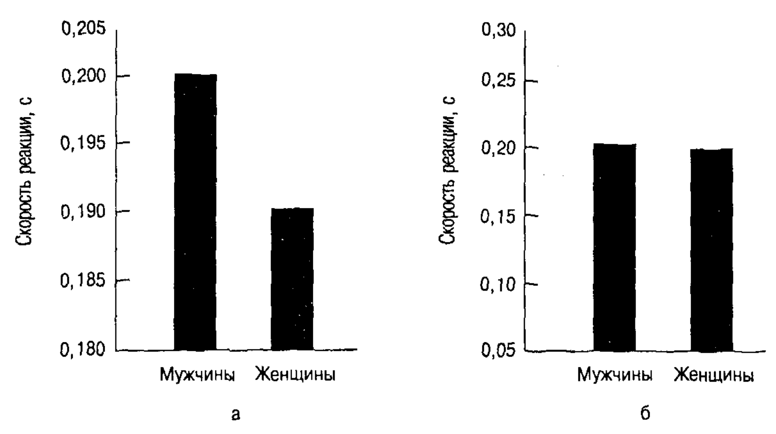


Рис. 4.9. Гистограмма, показывающая (а) повторение рис. 4.8, но с размеченной осью /и (б) те же данные, что и на рис. 4.8, но с иначе размеченной осью У

Мораль ясна: следите за осью У. Особенно скептично следует относиться к графикам, показывающим большое различие по оси У, при этом не размеченной, размеченной неясно или с очень малой ценой деления.

Статистика вывода

Как и большинству людей, исследователям необходимо, чтобы их любили и ими восхищались окружающие. Добиться этого можно, получив интересные результаты исследования, которые можно применить не только к данным, собранным в ходе экспериментов. Это означает, что хотя в исследовании рассматривается лишь небольшая *выборка* данных, которые возможно собрать, исследователи надеются, что сделанные ими выводы можно будет применить к более широкой *популяции.* Ведь задача исследовательской работы в целом и заключается в открытии общих законов поведения.

Чтобы проиллюстрировать применение статистики вывода, рассмотрим гипотетическое исследование, в котором сравнивается поведение в лабиринте двух групп крыс: одних кормят сразу после прохождения лабиринта, а других через 10 секунд. Эмпирический вопрос: способствует ли немедленное подкрепление запоминанию лабиринта? Предположим, что при исследовании 10 крыс (по 5 в каждой группе) получены следующие результаты. Каждая оценка означает количество попыток, потребовавшихся для запоминания лабиринта. Запоминание операционально определено как безошибочное прохождение лабиринта.

№ крысы	Немедленное кормление	№ крысы	Задержанное кормление
1	12	6	19
2	13	7	20
3	16	8	16
4	11	9	18
5	14	10	15

Обратите внимание, что оценки в каждой колонке не совсем одинаковы — это результат небольших различий между пятью крысами из каждой группы и, возможно, других случайных факторов. Однако, несмотря на отсутствие полного сходства, видно, что получавшие немедленное подкрепление крысы запомнили лабиринт быстрее (т. е. за меньшее количество попыток).

Но общего впечатления от чисел недостаточно. Первый шаг анализа — вычисление таких величин описательной статистики, как среднее арифметическое и стандартное отклонение:

	Немедленное подкрепление	**Отсроченное подкрепление**
Среднее арифметическое	**13,2**	**17,6**
Стандартное отклонение	**1,9**	**2,1**

В среднем, по крайней мере в этом примере, на запоминание лабиринта требуется больше попыток, если пищевое вознаграждение задерживается. Так же как показывает стандартное отклонение, изменчивость оценок в каждом наборе довольно низкая и практически одинакова для обеих групп. Можем ли мы заключить, что немедленное подкрепление *в целом* ускоряет запоминание лабиринта? Пока нет. Необходим заключительный анализ данных, в данном случае включающий проверку гипотезы[1].

Проверка гипотезы

Для проверки гипотезы сначала необходимо сделать предположение о том, что разница между двумя изучаемыми условиями не сказывается на выполнении заданий, в данном случае разница между немедленным и отсроченным вознаграждением. Это предположение называется **нулевой гипотезой** (нуль = ничто), обозначается H_0 (читается «аш нулевое»). Гипотеза исследования (меньшее количество попыток у крыс, получающих немедленное подкрепление), тот результат, который вы *надеетесь* получить, чтобы завоевать друзей и оказывать влияние на людей, называется **альтернативной гипотезой,** или H_1 Проводя исследование, вы будете стараться опровергнуть H_0 и соответственно подтвердить (но не доказать) H_1 — близкую вашему сердцу гипотезу.

Если эти слова звучат для вас странно, представьте, что дело происходит в суде. Обвиняемый человек считается невиновным, т. е. сделано предположение, что он ничего не совершал (нуль). Задача обвинения состоит в том, чтобы убедить суд в альтернативной гипотезе, а именно в том, что обвиняемый совершил преступление. Как и прокурор, исследователь должен показать, что нечто в действительности имеет место, а именно что в рассматриваемом случае задержка подкрепления влияет на запоминание лабиринта.

Заключительный анализ может привести к одному из двух результатов. Обнаруженные вами различия в поведении двух групп крыс могут быть вызваны настоящими, действительными и несомненными причинами или быть игрой случая. Это

[1] Вторая категория заключительного анализа называется «оценка». Она представляет собой оценивание показателей популяции на основании оценок отдельной выборки.

означает, что различия в пределах выборки могут отражать действительные явления, но могут и не делать этого, и соответственно результатом заключительного статистического анализа будет либо отказ от гипотезы H_0, либо ее подтверждение. Невозможность отвергнуть гипотезу H„ означает, что все найденные вами различия (а разница между группами почти всегда обнаруживается в ходе исследования), вероятнее всего, вызваны случайностью, ведь вам не удалось найти действительной закономерности, которую можно обобщить для случаев, выходящих за пределы выборки. Отказ от H_0 значит, что вы считаете, что некое явление действительно имело место в вашем исследовании и его результаты можно обобщить. В примере с лабиринтом отказаться от H_0 значит найти статистически значимые различия, свидетельствующие о существовании общей закономерности: немедленное подкрепление способствует запоминанию лабиринта.

Абсолютную истинность гипотезы исследования (H,) доказать невозможно, так же как нельзя *целиком и полностью* доказать виновность подсудимого: считается, что вина доказана тогда, когда не возникает обоснованных сомнений. Таким образом, H_0 может быть отвергнута (и в то же время может быть подтверждена H,) с определенной степенью уверенности, описываемой как значение альфа (а). Формально альфа означает вероятность получения определенных результатов, если H_0 истинна. Условно альфа принимается равной 0,05 ($\alpha = 0{,}05$), но можно задать также и другие значения альфа (например, $\alpha = 0{,}01$). Отказ от H_0 при значении альфа, равном 0,05, означает, что вы считаете вероятность того, что результаты вашего исследования вызваны случайными факторами, очень низкой (5 из 100). Если они не являются игрой случая, то должна быть какая-то другая причина, а именно (как вы надеетесь) изучаемое вами явление, в данном случае задержка подкрепления.

Выбор значения 0,05 связан с особенностями нормальной кривой, которые рассматривались выше. Вспомните, что для нормального распределения оценок вероятность того, что конкретная оценка будет лежать на расстоянии, превышающем два стандартных отклонения от среднего арифметического по оси *X,* довольно низкая — 5% или меньше. Такое случается редко. Аналогично при сравнении двух наборов оценок, как в случае с лабиринтом, нас интересует вероятность обнаружения различий между значениями среднего арифметического, если в действительности никаких различий не существует (т. е. если гипотеза H_0 истинна). Если вероятность достаточно низкая, мы отвергаем H_0 и считаем, что обнаружены действительные различия. «Достаточно низкая» — это вероятность, равная 5% ,или 0,05. Другими словами, если H_0 истинна, то обнаружение различий между значениями среднего арифметического настолько маловероятно (редкое событие), что мы просто не можем поверить, что H_0 истинна. Мы считаем, что произошло что-то другое (т. е. задержка подкрепления действительно снижает скорость запоминания лабиринта), а поэтому отвергаем H_0 и заключаем, что между группами существуют «статистически значимые» различия.

Ошибки 1 -го и 2-го рода

Из предыдущего примера ясно, что мы можем неверно решить, нужно отвергать или нет H_0. В действительности есть два вида таких ошибок. Во-первых, можно отвергнуть H_0 и считать подтвержденным II , радуясь тому, что сделано новое

значительное открытие, и ошибиться. Отказ от H_0, когда она по сути истинна, называется **ошибкой 1-го рода.** Вероятность такого события равна значению альфа, т. е. обычно 0,05. Таким образом, задание альфа значения 0,05 означает, что вероятность сделать ошибку 1-го рода, т. е. решить, что мы наблюдали некоторое явление и при этом ошибиться, — 5%. Ошибку 1-го рода можно заподозрить, если несколько раз подряд не удается получить нужные результаты при воспроизведении исследования.

Другой вид ошибки называется **ошибкой 2-го рода.** Она возникает, если вы не отвергаете H_0 и ошибаетесь. Это значит, что в ходе исследования вы не обнаружили то, что ожидали, расстроились из-за этого, но при этом ошиблись, поскольку в действительности данное явление наблюдается в популяции, но вы просто не смогли обнаружить его на изученной выборке. Ошибка 2-го рода нередко совершается, если проводимые измерения недостаточно надежны или точны для того, чтобы показать различия между выполнением задания при разных условиях. Как вы узнаете из главы 10, это часто случается в исследованиях по оценке программ. Проводимая программа может иметь значимое, но небольшое влияние на испытуемых, и поэтому измерения будут недостаточно тонкими, чтобы показать его.

В табл. 4.5 показано четыре возможных результата заключительного статистического анализа данных исследования, в котором сравниваются два экспериментальных условия. Несложно заметить, что правильное решение принимается, если H_0 отвергается, когда она ложна, и не отвергается, когда она истинна. Если H_0 ошибочно отвергается, то происходит ошибка 1-го рода, а если ошибочно не отвергается — ошибка 2-го рода. Если это поможет вам понять терминологию, в табл. 4.5 можно сделать следующие замены:

«H_0 не отвергнута» означает:

— вы провели исследование, сделали все необходимые анализы и получили ноль, ничто, отсутствие значимых различий. У вас действительно есть повод для расстройства, особенно если это ваш дипломный проект!

«H_0 отвергнута» означает:

— вы провели исследование, сделали все необходимые анализы и обнаружили значимые различия при значении альфа, равном 0,05. Жизнь наполнилась смыслом, вы удивите ваших друзей, а особенно вашего научного руководителя тем, что выполнили работу и *действительно что-то открыли]*

«H_0 истинна» означает:

— «несмотря на результаты вашего исследования, действительных различий не существует.

«H_0 ложна» означает:

— несмотря на результаты вашего исследования, действительные различия существуют.

Таблица 4.5

Получение статистического ответа: четыре возможных результата исследования со сравнением двух условий, *X* и *Y*

		Действительное положение дел	
		H_0 истинна: различий между *X* и *Y* не существует	**H_0 ложна:** различия между *X* и *Y* действительно существуют
Ваше статистическое решение	**H_0 не отвергнута:** в своем исследовании я не обнаружил значимых различий между *X* и *Y*, поэтому не отвергаю H_0	**Правильное решение**	**Ошибка 2-го рода**
	H_0 отвергнута: в своем исследовании я обнаружил значимые различия между *X* и *Y*, поэтому отвергаю H_0	**Ошибка 1-го рода**	**Правильное решение** (успех экспериментатора)

С учетом замен, названных выше, правильное решение означает, что а) действительных различий не существует и вы их не обнаружили или б) действительные различия существуют и вы их обнаружили (успех экспериментатора). Ошибка 1-го рода значит, что действительных различий нет, но на основании результатов исследования вы считаете, что они есть. Ошибка 2-го рода означает, что различия есть, но, проводя исследование, вы их не обнаружили.

Заключительный анализ

Принятие решения о том, отвергать ли нулевую гипотезу в исследовании, подобном эксперименту с научением в лабиринте, зависит от анализа двух основных видов изменчивости данных. Первый из них связан с различиями между «количеством попыток, требуемых для достижения критериальных» оценок для двух групп крыс. Эти различия вызываются комбинацией а) систематической дисперсии и (б) дисперсии ошибки. **Систематическая дисперсия** является результатом влияния конкретных факторов — либо изучаемой переменной (задержка подкрепления), либо факторов, которые не удается правильно проконтролировать[1]. **Дисперсия ошибки** — это несистематическая изменчивость, вызванная индивидуальными различиями между крысами в двух группах и любыми случайными, непредсказуемыми факторами, которые могут появиться в ходе исследования. Также индивидуальные различия и другие случайные факторы вызывают дисперсию ошибки в каждой группе и объясняют обнаруженные различия. Многие виды заключительного анализа данных включают вычисление следующего математического отношения:

[1] Эти неконтролируемые факторы называются «осложнителями». Более подробно они будут рассматриваться в следующей главе.

Статистика вывода = Изменчивость оценок при нескольких данных условиях (систематическая дисперсия + дисперсия ошибки)/Изменчивость оценок в пределах одного условия (дисперсия ошибки).

В идеале изменчивость, относящаяся к нескольким условиям, должна быть огромна, а изменчивость по отношению к одному условию относительно небольшой.

Очевидно, что именно такой результат получен для гипотетических данных из эксперимента с лабиринтом. Различия, обнаруженные между результатом двух разных условий, весьма основательны — крысам с отсроченным подкреплением требуется больше времени на изучение лабиринта (17,6 > 13,2), а оценки крыс в каждой группе довольно близки друг другу, что отражается небольшими значениями стандартного отклонения (2,1 и 1,9). В таком случае вы, вероятно, воспользуетесь особым видом анализа, называемым «проверка по критерию Стьюдента для независимых групп», — процедурой, известной вам из вводного курса статистики. В приложении *C* описано, как проводить такую проверку.

Мощность и сила эффекта

Все исследователи стремятся опровергнуть нулевую гипотезу, если она действительно ложна («успех экспериментатора» в табл. 4.5). Вероятность такого события описывается как **мощность** статистической проверки. На мощность влияет значение альфа, величина эффекта экспериментального воздействия, и в особенности величина выборки. Последняя характеристика находится под непосредственным контролем экспериментатора, и поэтому исследователи иногда проводят «анализ мощности» в самом начале исследования, чтобы определить подходящий размер выборки. Студенты нередко очень расстраиваются из-за того, что у них «не получается» исследование (т. е. не найдено значимых различий). Такой результат нередко является следствием небольшого размера выборки. В таком случае мощность невелика и, возможно, это спровоцирует появление ошибки 2-го рода, когда нечто действительно могло произойти в ходе исследования, но обнаружить это не удалось. В целом, ошибка 2-го рода и мощность обратно пропорциональны. Чем выше мощность, тем меньше вероятность, что будет сделана ошибка 2-го рода, и наоборот. С другой стороны, огромный размер выборки может привести к получению результата статистически значимого, но относительно бесполезного. Так, очень небольшие, но статистически значимые различия между группами могут почти не иметь практического значения для исследования с большим количеством участников[1]. Например, при достаточно большом количестве крыс в каждой группе можно провести исследование с лабиринтом, обнаружить разницу в скорости его прохождения, равную 1 секунде, и считать, что эта разница значима при альфе, равной 0,05. Однако различие это не будет иметь практически никакого значения. С другой стороны, если величина эффекта, который имеет экспериментальное воздей-

[1] Обратите внимание на слово «могут». Также возможно, что небольшие различия будут иметь большое практическое значение. Например, Розенталь и Росноу (Rosenthal and Rosnow, 1991) обнаружили взаимосвязь между приемом аспирина и возможностью избежать инфаркта. Влияние весьма слабое — прием аспирина снижает вероятность инфаркта приблизительно на 3%, но для группы риска, которая составляет около 1 миллиона людей, 3% означает снижение общего числа инфарктов в год на 30 000.

ствие, при небольшом размере выборки невелика, то значение этого факта может оказаться заметным.

Эта закономерность привела к тому, что исследователи стали уделять больше внимания **величине эффекта,** т. е. изменчивости зависимой переменной, чем независимой переменной (Cohen, 1988). Величина эффекта связана с относительной величиной различий между условиями эксперимента и размером выборки. В последние годы было разработано несколько «показателей величины эффекта», один из них *(d* по Коэну) представлен в приложении *С.* Важно отметить, что проверку гипотезы и определение величины эффекта можно рассматривать как взаимно дополнительные процедуры. Проверка гипотезы показывает, что экспериментальное воздействие имело определенный надежный эффект, а величина эффекта говорит о его величине. Другими словами, проверка гипотезы отвечает на вопрос, требующий ответа «да» или «нет», а определение величины эффекта — на вопрос «сколько?»

Познакомившись с основными инструментами работы с данными, вы готовы начать изучение трех первых глав об экспериментальном методе — самом мощном из орудий, помогающем понять сложные механизмы поведения и тонкости психических процессов. Мы начнем с введения в экспериментальный метод, рассмотрим проблемы, связанные с контролем за таким исследованием, и изучим особенности наиболее распространенных экспериментальных планов.

Резюме

Что измерять: разнообразие поведенческих актов

В психологии изучаются различные стороны поведения — от непосредственно наблюдаемых действий до результатов самоотчетов и показателей физиологической активности. Какие измерения проводить в ходе конкретного исследования, зависит от способа операционального определения конструктов, используемых в данном исследовании. Во многих областях психологии со временем были разработаны стандартные измерения (например, скорость реакции).

Оценочные измерения

Качественные измерения характеристик поведения являются одновременно надежными и валидными. Надежность означает возможность воспроизведения и малое количество ошибок измерений. Измерения валидны, если действительно измеряется то, что предполагается. Уверенность в валидности растет, если измерение разумно (очевидная валидность) и на его основании можно предсказать будущие события (критериальная валидность). Конструктной валидностью обладают исследования, в ходе которых изучаются отношения между измеряемым конструктом и связанным с ним явлением и которые дают стойкие, предсказуемые результаты.

Виды шкал измерений

Данные, собираемые в ходе психологических исследований, можно разделить на четыре группы в зависимости от используемой шкалы измерений: номинальной, порядковой, интервальной и шкалы отношений. В случае номинальной шкалы деления представляют собой ярлыки категорий, а изучается частотность оценок в

пределах каждой категории. Порядковая шкала используется, если изучаемые явления располагаются в ряд в определенном порядке. Интервальная шкала и шкала отношений имеют равные промежутки между делениями, но точку истинного нуля имеет только шкала отношений. Традиционно психологи предпочитают использовать интервальную шкалу и шкалу отношений, так как они позволяют применить большое количество видов статистического анализа.

«Кого» измерять: выборки

Если задача исследования состоит в изучении особенностей конкретной популяции, то сделанная для этого выборка должна быть репрезентативной и строиться как один из вариантов вероятностной выборки. При простой случайной выборке каждый член популяции имеет равные шансы быть выбранным. При расслоенной выборке важные подгруппы популяции представлены пропорционально. Кластерная выборка используется, когда нет возможности получить информацию о всех членах популяции. Большинство психологических исследований имеет своей целью не описание особенностей популяции, а выявление отношений между переменными, поэтому нередко наиболее подходящим является использование простой выборки. В большинстве психологических исследований используется удобная выборка, представляющая собой разновидность простой выборки.

Статистический анализ

В психологии статистический анализ является важнейшим инструментом, позволяющим понять значение результатов исследования. Описательная статистика применяется для обобщения результатов, полученных участниками конкретного исследования, и включает оценку общей тенденции (например, нахождение среднего арифметического, медианы и моды) и изменчивости (нахождение разброса, стандартного отклонения и дисперсии). Данные могут быть наглядно представлены в виде графиков (например, гистограмм или методом стебля и листа). Статистика вывода позволяет решить, являются ли результаты исследования случайностью или же они отражают действительную взаимосвязь явлений, которую можно распространить на более широкую популяцию. Целью заключительного анализа является отклонение гипотезы об отсутствии различий (т. е. нулевой гипотезы), если они в действительности существуют. Ошибка 1-го рода происходит, если нулевая гипотеза не должна быть отвергнута, но отвергается, а ошибка 2-го рода — когда экспериментальное воздействие дает определенный эффект, но в ходе исследования его не обнаруживают. Значение результатов исследования оценивают на основе информации о величине эффекта.

Задания для повторения

Выбор ответа

1. В XIX в. под влиянием экспериментов, посвященных психической хронометрии, было разработано учение о скорости реакции. В одном из таких исследований скорость простой сенсомоторной реакции (ПСР) составила

0,24 с, а скорость реакции выбора (РВ) — 0,29 с. Какой вывод был сделан о психическом явлении дифференциации?

1) Ее длительность составила 0,29 с;
2) она равна 0,29 + 0,24 с, или 0,53 с;
3) она составляет 0,29 - 0,24 с, или 0,05 с;
4) его длительность нельзя вычислить, так как СДР > СПР.

2. Какое из следующих утверждений о надежности измерений верно?
   1) измерение надежно, если оно представляется разумным по отношению к определенному конструкту (например, решение задач кажется разумным способом измерения уровня интеллекта);
   2) если измерение конструкта надежно, то при повторном измерении будут получены аналогичные результаты;
   3) чем больше ошибка измерений, тем выше надежность;
   4) результаты теста на надежность френологических измерений черепа будут низкими.
3. Исследователь провел измерения застенчивости и обнаружил, что студенты, получившие высокие оценки, также имеют высокий показатель интроверсии по известной шкале интроверсии—экстраверсии. Такие результаты *лучше всего* иллюстрируют:
   1) очевидную валидность;
   2) критериальную вал идность;
   3) надежность;
   4) конструктную валидность.
4. Каковы преимущества кластерной выборки по сравнению с простой случайной и расслоенной выборками?
   1) в отличие от остальных кластерная выборка является разновидностью вероятностной выборки;
   2) кластерная выборка не требует наличия полного списка популяции;
   3) кластерная выборка позволяет правильно отобразить подгруппы популяции (например, мужчин и женщин);
   4) только кластерную выборку можно назвать «удобной» выборкой.
5. Для какой из следующих шкал измерений приведен правильный пример?
   1) интервальная шкала — номера домов на улицах;
   2) шкала отношений — температура в градусах Цельсия;
   3) номинальная — как долго младенцы смотрят на новый стимул;
   4) порядковая — студенты колледжа решили, что профессор *X* самый строгий, профессор Y второй после него, и т. д.

Короткие эссе

1. В чем заключается суть исследования Дондерса, посвященного «психической хронометрии»? Каков его основной недостаток?

2. Дайте определение надежности и объясните, почему френологические измерения имели высокую надежность?
3. Опишите, как бы вы определили а) очевидную валидность и б) критериальную валидность теста.
4. На примере задержки вознаграждения покажите, как определяется конструктная валидность.
5. Охарактеризуйте основные различия между описательной статистикой и статистикой вывода.
6. Опишите различия между средним арифметическим, медианой и модой и покажите, в каких случаях медиана лучше описывает общую тенденцию, чем среднее арифметическое.
7. Опишите суть метода стебля и листа и объясните, почему он иногда больше подходит для отображения набора оценок, чем гистограмма.
8. В терминах шкал измерений объясните историю Дилберта, изображенную на рис. 4.4.
9. Опишите суть проверки гипотезы и объясните различия между ошибками 1-го и 2-го рода.

Упражнения

Упражнение 4.1. Выборки

1. С помощью таблицы случайных чисел, приведенной в приложении *D,* сделайте случайную выборку из вашей группы по экспериментальной психологии. Преподаватель предоставит вам список группы и задаст размер выборки.
2. Повторите данное упражнение, сделав расслоенную выборку, отражающую количество мужчин и женщин в вашей группе.

Упражнение 4.2. Шкалы измерений

Для каждого из приведенных ниже исследований определите, какая шкала измерений была использована при измерении характеристик поведения.

1. Салли хочет выяснить, по каким предметам дети республиканцев и дети демократов больше успевают, точным, гуманитарным или экономическим.
2. Фред решил исследовать, действительно ли крысы, изучившие один лабиринт, изучат второй быстрее, чем необученные.
3. Джим предполагает, что дети оценят цветные телевизионные программы выше, чем черно-белые, а у взрослых цвет не повлияет на оценку.
4. Нэнси считает, что соматотип изменяется с возрастом, и предлагает определять соматотипы у группы людей в 10, 15 и 20 лет по шкале Шелдона.
5. Сюзан изучает готовность людей помогать окружающим и считает, что она зависит от погоды — вероятность оказания помощи в солнечный день выше, чем в пасмурный.

6. Джон хочет узнать, какой из пяти новых сортов пива больше понравится (т. е. будет оценен как № 1) постоянным посетителям его бара.
7. Элен изучает, как студенты оценивают безопасность различных зданий студенческого городка. Она попросила нескольких студентов сложить карточки с написанными на них названиями зданий в стопку, в которой наиболее безопасные здания располагались бы сверху, а наименее безопасные — снизу.
8. Пэт считает, что люди с синдромом навязчивых состояний сделают меньше ошибок в составлении лабораторных отчетов по стандарту *APA,* чем здоровые люди.

Упражнение 4.3. H_0, H_1, ошибки 1-го и 2-го рода

Для каждого из следующих исследований а) определите нулевую гипотезу, б) сделайте предположение об альтернативной гипотезе, т. е. изложите возможный ход исследования, в) опишите результаты исследования при ошибке 1-го рода и г) охарактеризуйте результаты при ошибке 2-го рода.

1. В исследовании способности людей опознать ложь женщины и мужчины — участники исследования пытаются обнаружить обман в записанных на видео высказываниях женщин (в одних случаях они говорят правду, а в других — обманывают).
2. В исследовании восприятия младенцам дают привыкнуть к обычным изображениям человеческих лиц, а затем им показывают несколько неправильные лица, чтобы определить, видят ли они разницу.
3. Пациентов с депрессией и без нее просят высказать предположение о том, смогут ли они преодолеть лабиринт в человеческий рост.
4. Несколько спортсменов проходят тренинг формирования зрительных образов по новой методике непосредственно перед тем, как бить пенальти. Их результаты сравниваются с результатами других спортсменов, не проходивших тренинг.

Упражнение 4.4. Описательная статистика

Для гипотетического исследования по изучению лабиринта, в котором сравниваются немедленное и отсроченное вознаграждение, для каждой группы вычислите среднее арифметическое, медиану, разброс, дисперсию и стандартное отклонение.

Упражнение 4.5. Статистика вывода

Воспользуйтесь примером, приведенным в приложении *С,* и проведите проверку по критерию Стьюдента для описанного выше исследования по изучению лабиринта. Какие выводы вы сделаете? Будут ли крысы, получающие немедленное подкрепление, запоминать лабиринт быстрее?

ГЛАВА 5

Введение в экспериментальные исследования

Обзор задач главы

Четыре центральные главы этой книги, с 5 по 8, посвящены проблемам, возникающим при разработке экспериментов. В первой части главы 5 рассматриваются основные характеристики эксперимента — варьирование изучаемого фактора (независимой переменной), контроль за остальными факторами (внешние переменные) и измерение результата (зависимая переменная). Из второй части главы вы узнаете, как более или менее качественно осуществленная разработка исследования может влиять на его валидность. Изучив данную главу, вы:

- сможете дать определение управляемой независимой переменной и найти примеры ситуативных, рабочих и инструктивных переменных;
- будете понимать разницу между экспериментальной и контрольной группами;
- научитесь обнаруживать наличие в эксперименте осложняющих переменных и понимать, почему это осложнение затрудняет интерпретацию результатов эксперимента;
- будете различать независимые и зависимые переменные в описании эксперимента;
- будете различать управляемые и субъективные переменные и понимать смысл проблемы интерпретации, которая сопутствует использованию субъективных переменных;
- сможете распознавать факторы, которые могут уменьшить валидность статистических выводов и конструктную валидность эксперимента;
- сумеете описать различные варианты эффекта снижения внешней валидности эксперимента;
- сможете распознавать и описывать случаи, когда возникает угроза внутренней валидности эксперимента;
- будете понимать этические стороны работы в «фонде испытуемых».

Когда Роберт Вудвортс в 1938 г. опубликовал свою «Экспериментальную психологию», содержание этой книги было уже хорошо известно психологам. Еще в 1909 г. Вудвортс раздавал своим студентам в Колумбийском университете напечатанные на мимеографе копии брошюры «Проблемы и методы в психологии», а за-

тем, вместе с ней, «Руководство к лабораторным занятиям: эксперименты с памятью и т. д.», появившееся в 1912 г. К 1920 г. рукопись составила 285 страниц и была названа «Учебник по экспериментальной психологии». В 1932 г. напечатанный на мимеографе текст был переработан и наконец опубликован в 1938 г. К этому времени бывшие студенты Вудвортса уже обучали своих студентов по этому учебнику — он был так широко известен, что издатели объявили о нем просто: «Библия напечатана» (Winston, 1990).

Так называемая колумбийская библия была настоящей энциклопедией, содержала более 823 страниц основного текста и список литературы на 36 страницах. В главах, следующих за вводной, вся информация была разделена на 29 отдельных тем, таких как «Память», «Работа с лабиринтами», «Время реакции», «Ассоциации», «Слуховое восприятие», «Восприятие цвета» и «Мышление». Работая с книгой, студенты изучали методы, применяемые в каждой конкретной области, а также узнавали практически все, что было известно по каждой теме на 1938 г.

Влияние колумбийской библии на обучение экспериментальной психологии невозможно переоценить. Сегодняшнее обучение экспериментальной психологии и в некоторой степени даже структура книги, которую вы сейчас читаете, следуют образцу Вудвортса. В частности, он дал определение термину «эксперимент», до этого понимавшемуся широко, как практически любой вид эмпирического исследования (его определение мы используем и сегодня), и противопоставил эксперимент корреляционному исследованию (в наше время мы считаем такое разделение доказанным).

Отличительной особенностью экспериментального метода было управление явлением, названным Вудвортсом «независимой переменной», которое влияет на нечто, обозначенное им как «зависимая переменная». Как писал сам Вудвортс, экспериментатор «сохраняет постоянными все условия, за исключением одного, являющегося "экспериментальным фактором", или "независимой переменной". Наблюдаемое следствие представляет собой "зависимую переменную", которой в психологическом эксперименте является определенная характеристика поведения или эмоция, о которой сообщает испытуемый» (Woodworth, 1938, p. 2). Хотя эти термины не были изобретены Вудвортсом, он первый вложил в них тот смысл, который мы находим в них сегодня.

Тогда как в экспериментальном методе управляют независимыми переменными, при корреляционном исследовании, согласно Вудвортсу, «измеряется одна или несколько характеристик одного и того же человека [и] вычисляется корреляция между этими характеристиками. При использовании этого метода "независимые переменные" не используются и со всеми измеряемыми переменными обращаются одинаково» (Woodworth, 1938, p. 3). Более подробно корреляционные исследования рассматриваются в главе 9. Основное внимание этой и трех следующих глав сфокусировано на экспериментальном методе — наиболее мощном исследовательском инструменте выявления причинно-следственных связей.

Основные черты экспериментальных исследований

Со времен Вудвортса психологи рассматривают эксперимент как упорядоченное исследование, в ходе которого исследователь непосредственно изменяет некий фактор (или факторы), поддерживает остальные неизменными и наблюдает ре-

зультаты систематических изменений. Факторы, контролируемые в ходе эксперимента, называются независимыми переменными, факторы, поддерживаемые постоянными, — внешними переменными, а измеряемые поведенческие реакции — зависимыми переменными. Рассмотрим эти виды переменных более подробно.

Способы задания независимых переменных

Любой эксперимент можно охарактеризовать как исследование, в котором изучается влияние *X* на *Y*. *X* — это, по Вудвортсу, **независимая переменная.** Она представляет собой интересующий экспериментатора фактор, изучаемый с целью обнаружить его влияние на поведение. Иногда его называют «управляемый фактор», так как экспериментатор осуществляет над ним полный контроль и сам разрабатывает ситуации, с которыми встретятся участники в ходе исследования. Как вы увидите далее, понятие независимой переменной может быть расширено и распространено также на неуправляемые или субъектные переменные, но сейчас мы будем рассматривать только такие независимые переменные, которые полностью контролируются экспериментатором.

Независимые переменные должны принимать минимум два **значения,** т. е. в эксперименте должны сравниваться по крайней мере две ситуации (или два условия). Предположим, к примеру, что исследователь хочет узнать влияние различных доз марихуаны на время реакции. В таком исследовании должны сравниваться по крайней мере две различные дозировки. Будут определены «количество марихуаны» как независимая переменная, а также «дозировка № 1» и «дозировка № 2» как два значения независимой переменной. Разумеется, независимая переменная может принимать более двух значений; по сути, использование большего количества значений имеет заметные преимущества — об этом вы узнаете в главе 7, посвященной экспериментальным планам.

Экспериментальные исследования в зависимости от своих целей могут быть либо фундаментальными, либо прикладными, а проводиться они могут как в лаборатории, так и в полевых условиях (о различиях между этими видами опыта подробно рассказывается в главе 3). Эксперименты, проводящиеся в полевых условиях, иногда называются **полевыми экспериментами.** Термин **полевое исследование** — более широкий и касается любого эмпирического исследования за пределами лаборатории, включая экспериментальные исследования и исследования, проводящиеся неэкспериментальными методами.

Виды независимых переменных

Многообразие факторов, которые можно использовать в качестве независимых переменных, ограничено только пределами фантазии исследователя, но все независимые переменные, которыми управляют в ходе исследования, можно разделить на три пересекающиеся категории: ситуативные, рабочие и инструктивные переменные.

Ситуативные переменные — это различные особенности окружения, создаваемого для участников эксперимента. Например, если ученый, проводя исследование готовности людей оказывать помощь, хочет выявить влияние количества наблюдателей на вероятность оказания помощи, то он может создать ситуацию, в которой участники встречаются с человеком, которому нужна помощь. В одних случаях

участник один встречает человека, нуждающегося в помощи, а в других — участник и пострадавший могут быть окружены группой из трех или шести наблюдателей. В этом случае в качестве ситуативной независимой переменной может выступать количество людей, способных оказать помощь, не считая участника эксперимента, а ее значениями будут ноль, три и шесть (число наблюдателей).

Иногда экспериментаторы изменяют вид работы, выполняемой участниками. Один из способов управления **рабочими переменными** — давать для выполнения группам участников разные виды заданий. Например, в исследованиях по когнитивной психологии, чтобы определить наиболее распространенные виды ошибок, участников могут попросить решить несколько разных логических задач. Аналогично лабиринты могут различаться по уровню сложности, при исследовании восприятия могут использоваться разные виды иллюзий и т. д.

Инструктивными переменными можно управлять, попросив разные группы участников по-разному выполнить одно задание. Например, в ходе исследования памяти участникам, которым показывают один и тот же список слов, можно дать разные инструкции по способу запоминания. Одних можно попросить создавать визуальные образы для слов, других — устанавливать ассоциации между их смежными парами, а третьих — просто повторять каждое слово по три раза.

В одном исследовании можно также комбинировать различные виды независимых переменных. При исследовании влияния числа людей, их мотивации и сложности задания на способность к его решению одних участников можно поместить в большие, а других — в маленькие комнаты, тем самым воздействуя на фактор скопления людей с помощью ситуативной переменной размера комнаты. Далее, одним участникам в каждой комнате можно выдать сложные кроссворды, а другим — более легкие. Кроме того, с помощью инструктивной переменной можно воздействовать на мотивацию, сказав одним участникам, что они получат $1 за кроссворд, а другим — $5.

Контрольные группы

В некоторых экспериментах независимая переменная — это то, осуществляется или нет экспериментальное воздействие. В таком случае она принимает значения 1 или О — это означает, что одни участники испытывают воздействие, а другие — нет. Например, в исследовании влияния сцен насилия, показанных по телевизору, на детскую агрессивность одним детям могут показать передачу с жестокими сценами, а другим — не показывать. Термин **экспериментальная группа** используется для обозначения участников, испытывающих экспериментальное воздействие. Участники, не испытывающие воздействия, принадлежат к **контрольной группе.** В идеале, участники из контрольной группы абсолютно идентичны участникам из экспериментальной группы и отличаются от них только тем, что испытывают экспериментальное воздействие. Таким образом, контрольная группа дает базовые показатели, с которыми сравниваются оценки экспериментальной группы. Можно представить себе это следующим образом: контрольная группа = группа сравнения. В главе 7, первой из двух глав, посвященных экспериментальным планам, вы узнаете о различных видах контрольных групп.

Контроль за внешними переменными

Вторая особенность экспериментального метода состоит в том, что исследователь пытается контролировать так называемые **внешние переменные.** Они представляют собой неуправляемые факторы, сами по себе не интересующие исследователя, но способные повлиять на изучаемое поведение. Если их поддерживают постоянными, они не представляют опасности для исследования, но при отсутствии адекватного контроля над ними могут определенным образом воздействовать на исследуемое поведение. Такое явление называется осложнением. **Осложнитель** — это любая неуправляемая внешняя переменная, которая «ковариирует» с независимой переменной и может повлиять на интерпретацию результатов. Осложняющие переменные сами по себе изменяют то, что изменяют независимые переменные (т. е. эти переменные «ковариируют»), и поэтому их воздействие невозможно различить. Таким образом, если в исследовании имеется осложнитель, результаты такого исследования могут быть вызваны как осложняющей переменной, так и независимой, а также их комбинацией. Узнать при этом, чем в действительности вызываются результаты, не представляется возможным.

Чтобы проиллюстрировать действие простого осложнителя, рассмотрим один эксперимент. Исследователь поставил перед собой задачу доказать, что студенты, пытающиеся сразу изучить большое количество материала, не имеют таких хороших результатов, как те, кто разделяет учебу на несколько частей. Другими словами, быстрое зазубривание предполагается менее эффективным, чем равномерно распределенное изучение материала. Выбираются три группы студентов и каждой их них дается задание: выучить пять глав учебника по общей психологии. Первой группе дают на изучение материала три часа в понедельник, второй — три часа в понедельник и три часа во вторник, а третьей — по три часа в понедельник, вторник и среду. В пятницу проверяется знание материала участниками всех трех групп (план эксперимента см. в табл. 5.1). Результаты показывают, что участники группы 3 получили самые высокие оценки, более низкие оценки получили участники группы 2, а группа 1 практически не справилась с заданием. Исследователь заключает, что распределенные занятия более эффективны, чем массированные. Согласны ли вы с таким выводом?

Таблица 5.1

Осложнения в гипотетическом эксперименте по распределению нагрузки

	Понедельник	Вторник	Среда	Четверг	Пятница
Группа 1	3	—	—	-	Экзамен
Группа 2	3	3	-	-	Экзамен
Группа 3	3	3	3	-	Экзамен

Примечание. Число 3 в каждой клетке равняется количеству часов, использованных для изучения пяти глав учебника по общей психологии.

Скорее всего, вы не согласитесь с выводом экспериментатора, поскольку в этом исследовании можно легко обнаружить по крайней мере два осложнителя. Распределение занятий у участников действительно различается (1, 2 и 3 дня), но **также**

различается и общее время занятий (3,6 и 9 часов). Это прекрасный пример осложнителя, так как невозможно определить, вызваны ли результаты одним фактором (распределение нагрузки) или другим (общее время занятий), ведь два этих фактора отлично ковариируют. Для описания этой ситуации можно сказать, что «распределение нагрузки осложнено общим временем обучения». Второй осложнитель не так очевиден, но вызывает не меньше проблем. Речь идет о времени, проходящем до проверки результата. Проверка всех групп проводится в пятницу, но время, прошедшее с момента изучения материала, у каждой группы разное. Возможно, группа 3 лучше всего справилась с заданием потому, что ее участники изучали материал позднее и меньше успели забыть. Таким образом, в этом эксперименте фактор распределения нагрузки осложнен одновременно общим временем обучения и временем, прошедшим до проверки. Каждый осложнитель сам по себе может повлиять на результаты, но также они могут воздействовать совместно и изменить общую картину каким-то особым образом.

Взгляните на табл. 5.2. С ее помощью можно легко выявить осложнители. В первом столбце приведены значения независимой переменной, а в последнем — результаты проверки. В двух средних столбцах даны внешние переменные — их необходимо поддерживать постоянными, используя соответствующие стратегии (см. следующую главу). Если они не остаются постоянными (как в данном случае), то возникает осложнение. Очевидно, что результаты можно объяснить различиями в любом из первых трех столбцов, как взятых по отдельности, так и в различных комбинациях. Чтобы исправить ситуацию с осложнителями в данном примере, необходимо добиться, чтобы значения в двух средних столбцах не изменялись.

Таблица 5.2

Обнаружение осложнителей

Значения НП — Распределение нагрузки, дни	ВШ Время занятий, часы	ВП 2 Время до проверки, дни	ЗП Результат проверочного теста
1	3	3	Низкий
2	6	2	Средний
3	9	1	Высокий

НП — Независимая переменная.
ВП — Внешняя переменная.
ЗП — Зависимая переменная.

При работе над упражнениями в конце главы вам придется находить осложнители. Представление условия в виде таблицы, подобной табл. 5.2, облегчит вам выполнение задания. Попробуйте усовершенствовать эксперимент с распределением нагрузки. Как отделить осложнители от внешних переменных?

Научиться помнить о возможности осложняющих факторов и разрабатывать механизмы контроля за ними — значит сформировать один из наиболее сложных навыков, необходимых для научного мышления. Далеко не все осложнители так очевидны, как в примере с массированной/распределенной нагрузкой. В остальных главах мы будем часто встречаться с этой проблемой и кратко рассмотрим ее при обсуждении так называемой внутренней валидности исследования.

Измерение зависимых переменных

Третий этап эксперимента — измерение характеристики поведения, предположительно испытавшей воздействие независимой переменной. Для описания характеристик поведения, которые измеряются в ходе эксперимента и являются его результатами, используется термин **зависимая переменная.** Если, как отмечалось ранее, эксперимент можно описать как воздействие X на Y *и X* — независимая переменная, то *Y* — это зависимая переменная. В исследовании влияния телевизионных сцен насилия на детскую агрессивность зависимой переменной является агрессивность. В исследовании фактора распределения нагрузки — результаты экзамена.

Правдоподобность любого эксперимента и возможность сделать существенное открытие отчасти определяются правильностью выбора зависимой переменной, измеряемой в ходе исследования. Как мы уже знаем, на эмпирический вопрос можно получить ответ, только если вопрос сформулирован с определенной степенью точности. Вы можете потратить немного времени и еще раз просмотреть главу о постановке эмпирических вопросов — она напомнит вам о понятии операционального определения. Один из ключевых моментов разработки эксперимента касается операциональных определений характеристик поведения, выбранных и измеряемых как независимые переменные. Воспроизведение эксперимента возможно, только если отдельные характеристики поведения точно определены.

В заключение отметим: очень важно понимать, что в зависимости от задач исследования конструкт может представлять собой независимую, внешнюю или зависимую переменную. В ходе эксперимента конструктом можно управлять как независимой переменной, пытаться контролировать его как внешний фактор или измерять как зависимую переменную. Рассмотрим, к примеру, конструкт «тревожность». Его можно представить в виде управляемой независимой переменной, если сказать участникам, что они будут подвергаться слабому, или напротив, болезненному, электрическому шоку, а затем спросить их, как они предпочитают ждать, в одиночестве или вместе с другими испытуемыми. Тревожность также может выступать как фактор, который в ходе эксперимента необходимо поддерживать постоянным. Например, если требуется оценить влияние семинара, посвященного публичным выступлениям, на умение студентов произносить речи, то не следует производить видеозапись студентов в одной группе, не записав также и другую группу. Если произведена видеозапись всех студентов, то уровень тревожности, вызванной фактором съемки, будет постоянным для всех групп. Кроме того, тревожность может быть зависимой переменной, как, например, в исследовании влияния различных видов экзаменов (например, заданий на выбор ответа или на написание эссе) на тревожность студентов во время сессии. В этом случае можно произвести особое психологическое измерение тревожности. Тревожность можно рассматривать и как характеристику личности, подразумевая, что одни люди обладают ею в более высокой степени, чем другие. Последнюю возможность мы рассмотрим в следующем разделе.

Управляемые и субъективные переменные

До сих пор термин «независимая переменная» означал некий фактор, непосредственно изменяемый исследователем. В таком случае в ходе эксперимента сравниваются условия, созданные экспериментатором и находящиеся под его контролем, с другими условиями. Но во многих исследованиях сравниваются группы людей, которые отличаются друг от друга иначе, чем это было запланировано при разработке исследования. Тогда сравнение производится между факторами, обозначаемыми как переменные *ex post facto*[1], переменные стихийно возникших групп, неуправляемые, или субъективные (дополнительные), переменные. Я предпочитаю использовать последний термин, но все они относятся к характеристикам участников исследования, таким как пол, возраст, интеллект, психические или физические нарушения или любые другие свойства личности, которые вы только сможете назвать. При использовании в исследовании субъектных переменных экспериментатор не имеет возможности управлять ими непосредственно, поэтому он *отбирает* людей для эксперимента на основании присущих им свойств.

Для иллюстрации различий между управляемыми и субъективными переменными рассмотрим гипотетическое исследование влияния тревожности на способность людей проходить лабиринт. Тревожностью можно *управлять* непосредственно, создав ситуацию, в которой одна группа будет испытывать тревожность (например, можно сказать, что участники будут выполнять задание перед лицом большого количества зрителей), а вторая — не будет (зрители будут отсутствовать). Человек, согласившийся принять участие в таком исследовании, является потенциальным участником любой из этих групп. С другой стороны, чтобы провести исследование с использованием *субъективной* переменной, можно подобрать две группы, различающиеся по уровню присущей им тревожности, и предложить участникам пройти лабиринт. В первую группу войдут люди с высокой тревожностью (предварительно выявленной с помощью специального теста). Вторая группа будет включать более спокойных людей. Обратите внимание, что существует заметная разница между таким исследованием и экспериментом с управляемой переменной. Если тревожность выступает как субъективная нереременная, люди, желающие принять участие в исследовании, не могут произвольно помещаться в любую из двух групп (всегда чего-то боящегося Фреда нельзя помещать в группу с низкой тревожностью), но в зависимости от свойств, которые сформировались у них *еще до* участия в исследовании, должны быть помещены только в одну из них.

Одни исследователи в соответствии с первоначальным значением термина «независимая переменная», примененного Вудвортсом, предпочитают использовать его для переменных, непосредственно управляемых экспериментатором. Другие рассматривают субъектные переменные как частный случай независимых переменных на том основании, что экспериментатор обладает некоторой степенью контроля над ними за счет предварительного отбора. Я придерживаюсь второй точки зрения и буду использовать термин «независимая переменная» в более широком смысле. Но если вы поймете разницу между управляемыми и неуправляемыми, или

[1] Имеющие обратную силу *(лат.)*.

субъективными, переменными, то для вас будет не так важно, используется ли термин «независимая переменная» в широком смысле (управляемая + субъективная) или в более узком (только управляемая).

Пример 4. Использование субъективных переменных

Эксперимент, проведенный в 1992 г. Шнейдером и Бьерклендом, является хорошим примером одновременного исследования нескольких субъективных переменных. Также он иллюстрирует процесс воспроизведения и дополнения исследований (см. главу 3). Ученые заинтересовались вопросом, как воздействуют на память: а) специальные знания в определенной области и б) общие способности. Предыдущие исследования концентрировались только на знаниях, поэтому важным дополнением, вносимым их исследованием, было «рассмотрение уровня общих способностей детей» (р. 462). Таким образом, в исследовании использовались две субъективные переменные[1]. Первой был объем знаний о футболе. Эта переменная принимала два значения: знаток и новичок. Второй переменной были общие способности, определяемые по «тесту когнитивных способностей». Как это часто бывает при использовании субъективных переменных, исследование состояло из двух частей: первая заключалась в оценке субъективных переменных для формирования отдельных групп, а вторая — в проведении самого эксперимента. В первой части исследования большая группа учеников второго и четвертого классов прошла тест на когнитивные способности и тест, по которому определялись познания в области футбола. По полученным при тестировании результатам дети были отобраны для участия во второй части исследования и разделены на четыре группы: знатоки футбола с высокими способностями, знатоки футбола с низкими способностями, новички в футболе с высокими способностями и новички в футболе с низкими способностями.

Во второй части исследования дети выполняли задание на классификацию и запоминание: им раздали наборы терминов и попросили разделить их на группы, а сразу после этого — запомнить их. Задание выполнялось дважды, один раз на примере слов, связанных с футболом, а второй раз на примере других слов. (Одна половина детей выполняла задания именно в таком порядке, а другая — в обратном. Этот метод называется позиционным уравниванием и будет обсуждаться в следующей главе.) Результат получался сложный, но главным было то, что дети с высокими способностями в целом более успешно выполнили оба вида заданий, чем дети с низкими способностями вне зависимости от их уровня футбольных знаний. Однако при учете исключительно познаний в области футбола результаты показывали, что знатоки футбола лучше справились с заданием, посвященным футболу. Так, при выполнении задания с футбольными терминами знатоки запомнили 62% информации, а новички — только 44%. Но при выполнении задания со словами, не связанными с футболом, обе группы показали примерно одинаковые оценки: 52% для знатоков футбола и 48% для новичков.

Только исследование, в котором используются *управляемые* независимые переменные, можно назвать экспериментом в самом строгом смысле этого слова. Иногда

[1] В действительности кроме этих было еще три: пол, интерес к футболу и год обучения, но мы сконцентрируемся на двух важнейших.

такой эксперимент называют «истинным» (что звучит несколько претенциозно и приводит к мысли, что все остальные исследования «ложные»). Исследования, использующие независимые субъективные переменные иногда называют исследованиями *ex post facto,* или квазиэкспериментами («квази» в данном случае означает «в определенной степени»)[1]. Иногда (по сути, довольно часто) исследования имеют дело и с управляемыми, и с субъективными независимыми переменными. Очень важно осознавать наличие субъективных переменных, так как они влияют на выводы, которые делаются на основании результатов исследования.

Как делать выводы, если используются субъективные переменные?

Поставьте галочку напротив этого раздела — он чрезвычайно важен. Вспомните: в главе 1 отмечалось, что одной из целей психологических исследований является объяснение поведения. Это означает, что мы хотим узнать, чем вызвано определенное поведение. Проще говоря, если используются управляемые переменные, то выводы о причинах поведения могут быть сделаны, а если субъективные — не могут. Однако суть дела состоит в степени контроля, который осуществляет экспериментатор в каждом из этих случаев.

В случае управляемых переменных эксперимент соответствует описанным в главе 1 условиям выявления причинно-следственной связи. Независимые переменные предшествуют зависимым, ковариируют с ними и в случае, если нет осложнителей, принимаются как наиболее разумное объяснение результатов. Другими словами, если вы изменяете один фактор и успешно поддерживаете остальные постоянными, то результаты могут объясняться *только* изменяемым фактором. В экспериментальном исследовании с двумя группами участников, не содержащем осложнителей, группы будут равноценными (т. е. любые различия будут случайными) по всем показателям, кроме управляемого фактора.

При использовании субъективных переменных экспериментатор также может варьировать определенный фактор (т. е. отбирать участников, имеющих конкретные характеристики), но не может поддерживать все остальные факторы постоянными. Отбор участников по степени предрасположенности к тревожности (высокой или низкой) не гарантирует, что две группы будут эквивалентны по другим показателям. Фактически они могут различаться по многим показателям (например, по степени уверенности в себе), которые способны повлиять на результаты исследования. Если в ходе такого исследования возникает различие между группами, то мы не можем сказать, что его *причиной* является субъективная переменная. В аспекте условий причинно-следственной связи это значит, что если независимая переменная предшествует зависимой и ковариирует с ней, то нельзя исключить возможность альтернативного объяснения их взаимосвязи, поскольку отсутствует контроль за некоторыми внешними факторами. При наличии субъективных переменных мы можем говорить лишь о том, что результаты групп различаются по зависимому показателю.

[1] Термин «квази-экспериментальный план» по сути является более широким обозначением любого плана, в котором участники не могут быть случайным образом разделены на группы, которые в дальнейшем будут изучаться (Cook & Campbell, 1979). Такие планы нередко встречаются в прикладных исследованиях и рассматриваются в главе 10.

Пример из социальной психологии поможет понять разницу. Предположим, вы интересуетесь альтруистическим поведением и хотите выяснить его зависимость от самооценки. Исследование можно провести двумя способами. Во-первых, можно непосредственно управлять самооценкой. Для этого участникам предварительно раздают тест личности, а затем сообщают им ложную информацию о результатах тестирования. В зависимости от того, какая информация сообщается, позитивная или негативная, самооценка может быть временно поднята или снижена. Далее можно попросить участников выполнить по желанию какую-либо работу и посмотреть, будут ли люди с высокой самооценкой помогать охотнее. Второй способ проведения исследования заключается в том, чтобы раздать участникам надежный и валидный тест личности для определения уровня самооценки, а затем разделить их на две группы: в одну войдут люди, чьи показатели выше 25%, а в другую — ниже этого уровня. Самооценка в этом случае является субъективной переменной, одна половина участников будет иметь низкую самооценку, а другая — высокую. Как и в первом исследовании, участников из этих двух групп можно попросить о добровольном участии в некоторой деятельности.

В первом исследовании различия в желании выполнять работу могут быть *непосредственно* привязаны к изменениям самооценки. Если все остальные факторы должным образом контролируются, различия в желании оказывать помощь могут быть вызваны *только* временным снижением или повышением самооценки. Во втором исследовании, однако, нельзя сказать, что высокая самооценка является непосредственной причиной желания помогать. Можно сказать лишь, что вероятность оказания помощи людьми с высокой самооценкой выше, чем людьми с низкой самооценкой. Все, что вы можете, — это строить предположения о причинах этого явления, так как участники могут отличаться друг от друга по каким-то другим, неизвестным вам показателям. Например, люди с высокой самооценкой могли уже иметь опыт добровольной помощи, который мог поднять или укрепить их самооценку и увеличить вероятность оказания ими помощи в будущем. Или они могут обладать специальными навыками, способствующими их добровольному участию в данного рода работе (например, навык публичных выступлений). Как вы узнаете из главы 9, в корреляционных исследованиях при попытках сделать выводы проблемы возникают именно из-за трудности интерпретации результатов.

Вернемся ненадолго к исследованию Шнейдера и Бьеркленда (Schneider and Bjorkland, 1992), в котором в качестве субъективных переменных использовались знания о футболе и общие способности. Авторы исследования тщательно избегали выводов о причинно-следственной связи. Слово «причина» никогда не встречается в их статьях, а описание результатов всегда приводится в следующем виде: «оценки у одной группы выше, чем у другой».

Прежде чем переходить к дальнейшему обсуждению валидности экспериментальных исследований, прочитайте вставку 5.1. В ней характеризуются переменные, которые были использованы в одном из экспериментов с куклами бобо (классическом исследовании, знакомом вам из курса по общей психологии), первом исследовании подражательной агрессии. Изучение данного примера позволит вам применить знания о независимых, внешних, и зависимых переменных и увидеть, как управляемые и субъективные переменные могут встретиться в одном исследовании.

ВСТАВКА 5.1

Классические исследования— куклы бобо и агрессия

Спросите любого студента, только что прослушавшего курс по детской психологии, социальной психологии или психологии личности (а возможно, даже и по общей психологии) об экспериментах с куклами бобо, и он сразу поймет, о чем речь, и расскажет вам что-то вроде «Да, я знаю, это эксперименты, которые показали, что дети бьют надувных кукол, если видят, как это делают взрослые». Описание одного из экспериментов поможет нам понять разницу между независимыми, внешними и зависимыми переменными. Результаты исследования были опубликованы Альбертом Бандурой и его коллегами в 1963 г. под названием «Имитация моделей агрессивности, представленных посредством видеозаписи» *[Imitation of Film-Mediated Aggressive Models,* Bandura, Ross, &Ross, 1963).

Задание независимых переменных

Эксперимент включал как управляемые так и субъективные переменные. Основной управляемой переменной был вид опыта, предшествующего возможной агрессии. Переменная принимала четыре значения, которым соответствовали три экспериментальные группы и одна контрольная.

Экспериментальная группа 1: агрессия в реальной жизни (дети непосредственно наблюдали модель агрессивного поведения взрослых с куклой бобо).

Экспериментальная группа!: агрессия, воспринимаемая в видеозаписи (дети просматривали модели агрессивного поведения взрослых с куклой бобо в видеозаписи).

Экспериментальная группа?,: агрессия, воспринимаемая в мультфильмах (дети просматривали мультфильм «Кот Герман» с агрессией, направленной против мультипликационной куклы бобо).

Контрольная группа: модель агрессивного поведения не демонстрировалась.

Неуправляемой независимой переменной (субъективной переменной) являлся пол. В исследовании участвовали дети мужского и женского пола из детского сада при Стэнфордском университете (средний возраст - 52 месяца). (В действительности была еще одна управляемая переменная: участникам из групп 1 и 2 показывали модель с участием людей либо того же, либо противоположного пола). Базовая экспериментальная процедура заключалась в том, что детям показывали определенный вид агрессивности (или не показывали, в случае контрольной группы), а затем помещали их в комнату, заполненную игрушками (в числе прочих - кукла бобо), тем самым предоставляя возможность самим проявить агрессивность.

Контроль за внешними переменными

Были тщательно устранены некоторые возможные осложнители. Во-первых, для групп 1 и 2 взрослые демонстрировали агрессию, направленную против 5-футовой куклы бобо, а когда детям дали возможность самим поколотить куклу, их поместили в комнату с 3-футовой бобо. Тем самым отношение размеров куклы и человека поддерживалось приблизительно постоянным. Во-вторых, участников во всех четырех группах слегка расстроили перед тем как дали возможность проявить агрессию. Им разрешили немного поиграть с очень привлекательными игрушками, а затем экспериментатор сказал, что эти игрушки не простые и предназначены для других детей. Таким образом *все* дети испытали приблизительно одинаковое нарастание уровня эмоционального возбуждения до того, как им дали возможность проявить агрессию. Поэтому любые различия в агрессивности должны были быть связаны с подражанием, а не с эмоциональными различиями между группами.

Измерение зависимых переменных

В этом эксперименте использовались несколько различных способов измерения агрессивности. Агрессивные реакции классифицировались как подражательные, частично подражательные или неподражательные в зависимости от того, насколько близки они были к соответствующей модели поведения. Например, в операциональное определение подражательного агрессивного поведения входили удары куклы деревянным молотком, удары кулаком в нос и всевозможные пинки, частично подражательного - удары молотком по другим предметам и сидение на кукле без нанесения ей ударов, а неподражательная агрессия включала стрельбу дротиками из ружья в другие мишени, кроме куклы бобо, а также агрессивные действия, направленные против других объектов в комнате.

Если кратко сформулировать результаты, то дети в группах 1, 2 и 3 показали значительно более высокую степень агрессии, чем дети в контрольной группе, но при этом степень агрессии была одинаковой при различных моделях поведения. Кроме того, мальчики при любых условиях вели себя более агрессивно, чем девочки. Также некоторые связанные с полом различия наблюдались в особенностях проявления агрессии: девочки «охотнее сидели на кукле бобо, но (в противоположность мальчикам) воздерживались от нанесения ударов» (Bandura et al., 1963, p. 9). Рисунок 5.1 обобщает результаты исследования.

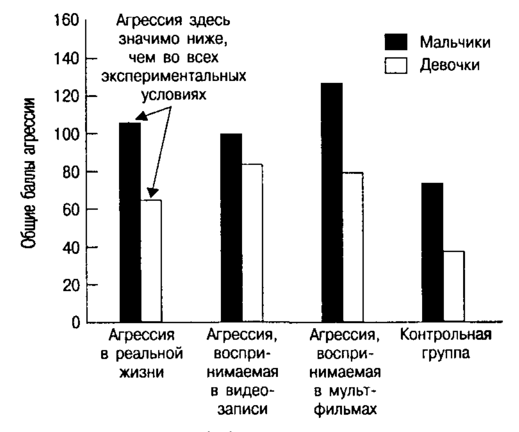


Рис. 5.1. Данные исследования с куклами бобо, проведенного Бандурой, Россом и Россом в 1963 г. для изучения влияния подражания на агрессию

Валидность экспериментальных исследований

В главе 4 было введено понятие валидности в применении к измерениям. Этот термин также применяется к эксперименту в целом. Так же как измерение считается валидным, если измеряется именно то, что предполагается, психологическое исследование считается валидным, если оно дает такое объяснение поведения, которое, как предполагается, оно должно давать. Эта часть главы познакомит вас с четырьмя различными видами валидности по схеме, предложенной Куком и Кэмпбеллом (Cook and Campbell, 1979) для полевых исследований, но применимой также и для любых других психологических исследований. Эти четыре вида — валидность статистических выводов, конструктная валидность, внешняя и внутренняя валидность.

Валидность статистических выводов

В предыдущей главе вы познакомились с применением статистики в психологии. В частности, вы узнали о видах шкалы измерений, проверке гипотез и получили представление о различиях между описательной статистикой и статистикой вывода. **Валидность статистических выводов** связана с границами корректного использования статистики и получением на основании статистического анализа правильных выводов.

Статистическая валидность исследования может снижаться по нескольким причинам. Во-первых, исследователь может неверно проводить анализ данных или нарушать условия, которые необходимо соблюдать при проведении конкретного вида анализа. Например, данные исследования могут быть расположены на порядковой шкале, а следовательно, потребуется использование определенной статистической процедуры. Исследователь, однако, по ошибке может использовать вид анализа, подходящий только для данных, представленных с помощью интервальной шкалы или шкалы отношений. Во-вторых, исследователь может избирательно сообщить об одних результатах анализа, полученных в соответствии со сделанными предположениями, и скрыть остальные (догадайтесь, какие) — такая практика граничит с мошенничеством (см. главу 2). Третий фактор, который способен снизить статистическую валидность исследования, — это надежность проводимых измерений. Если зависимые показатели ненадежны, возникает множество ошибок, что снижает вероятность обнаружения значимого влияния одних показателей на другие. Если такое влияние действительно существует (т. е. нулевая гипотеза должна быть отвергнута), а низкая надежность не позволяет его обнаружить, то результатом будет ошибка 2-го рода по вышеприведенной классификации.

Грамотные исследователи проводят планирование статистического анализа одновременно с разработкой экспериментального плана. По сути, эксперимент не может быть разработан без выяснения способов анализа данных.

Конструктная валидность

В предыдущей главе конструктная валидность описывалась в связи с измерением психологических конструктов: она свидетельствует одновременно о том, действительно ли в ходе тестирования измеряется некий гипотетический конструкт (например, задержка удовольствия) и соответствует ли этот конструкт экспериментальной реальности. В случае экспериментальных исследований термин **конструктная валидность** имеет сходное, но несколько отличающееся значение: он обозначает, насколько адекватно определяются независимые и зависимые переменные, используемые в исследовании. В случае исследования влияния показа по телевидению жестоких сцен на детскую агрессивность для определения конструктной валидности можно задать следующие вопросы: а) отражают ли наилучшим образом отобранные экспериментатором программы контраст между программами, содержащими сцены жестокости, и без них, и б) является ли они наилучшим из возможных способ измерения агрессии? Если в исследовании сравниваются жестокие мультипликационные персонажи (например, Элмер Фадд, стреляющий в Багса Банни) с обычными персонажами, (например, Винни-Пух), то на это могут возразить, сказав, что фантастические образы никак не могут влиять на агрессив-

ное поведение детей, поэтому более валидный способ управления независимой переменной, называемой «уровень жестокости, представленной в видеозаписи», будет заключаться в показе детям реальных видеозаписей людей, демонстрирующих разную степень жестокости.

Аналогично критике может быть подвергнут применяемый в конкретном исследовании способ измерения агрессии — и это действительно вызвало проблемы при исследовании данного явления. По вполне очевидным этическим причинам вы не должны разрабатывать исследование, в котором участники будут колотить друг друга. Агрессия определялась операционально самыми различными способами, и некоторые из них могут показаться вам более валидными (например, предварительно разозленные испытуемые считали, что наказывают ударом тока других испытуемых), чем другие (например, гудение автомобилей раздраженных водителей). Так же как в случае конструктной валидности, которая обсуждалась в предыдущей главе, посвященной валидности измерений, валидность определения независимых и зависимых переменных вырабатывается по мере того, как в ходе исследования накапливаются данные и появляется возможность их описания при помощи последовательной (и сводимой к единому основанию) системы.

Внешняя валидность

Психологов-экспериментаторов часто критиковали за то, что о студентах-второкурсниках и белых крысах они знают очень много, а обо всем остальном — очень мало. По сути, это критика **внешней валидности** — возможности генерализации данных, полученных в ходе исследования, за пределами содержания конкретных экспериментов. Чтобы исследование обладало наивысшей степенью внешней валидности, должна существовать возможность генерализации его результатов по трем направлениям: на другие популяции, другие условия и другое время.

Другие популяции

Именно к этому разделу относится замечание насчет крыс и второкурсников. Как мы узнали из главы 2, часть споров о правомерности исследований с животными касается того, насколько важны открытия, сделанные в ходе таких исследований, для объяснения человеческого поведения. Что касается второкурсников, вспомните, что Милграм сознательно не хотел использовать студентов и в качестве объектов для исследования подчинения и отобрал взрослых людей из нейтральной в этом отношении популяции. Однако, о большинстве психологов нельзя сказать того же. Обзор, проведенный Сирсом (Sears, 1986) в ходе исследования по социальной психологии, показал, что 75% исследований, опубликованных в этой области, в 1980 г. проводились с использованием студентов. Повторный обзор для исследований, опубликованных в 1985 г., дал результат 74%. Сире утверждал, что особенности студентов колледжей как популяции легко могут исказить общие выводы о социальных явлениях. В сравнении с основной популяцией, к примеру, студенты колледжей обладают более высокими когнитивными способностями, они более эгоистичны и восприимчивы к социальным воздействиям, а также легче меняют свое отношение к различным вопросам. По Сирсу, если в исследовании изучаются вопросы, связанные с этими особенностями, результаты, полученные при участии

студентов, могут не быть пригодными для других групп людей. Ученый считал, что исследователи должны расширить свои базы данных и воспроизводить важные открытия на различных популяциях. Но он также указал, что во многих областях исследований (например, при изучении восприятия) данные особенности студентов практически не влияют на получаемые результаты и, кроме того, студентов очень много, и поэтому они являются классическим примером удобной выборки (глава 4). Как показано во вставке 5.2, к этой группе применяется особый этический подход.

ВСТАВКА 5.2

Этика — набор участников - все становятся членами фонда

Большинство психологов-исследователей работают в колледжах и университетах и поэтому окружены потенциальными «морскими свинками» в виде студентов. Поскольку студенты не всегда охотно принимают участие в исследованиях, многие психологические факультеты университетов организуют так называемые «фонды испытуемых», или, как это называется в последнее время, «фонды участников». Этот термин обозначает группу студентов, обычно посещающих вводный курс по психологии, которых просят участвовать в исследованиях как в обязательных для данного курса занятиях. Если вы студент большого университета, то, возможно, уже имели опыт «добровольного» участия в двух или трех экспериментах, чтобы избежать снижения оценки за весь курс. Если в большом университете курс по общей психологии каждый семестр посещают 800 студентов и каждый студент записывается для участия в трех экспериментах, то исследователи получают возможность работать с 2400 участниками.

Фонды испытуемых облегчают исследователям поиск участников, а их существование оправдывают тем, что участие в исследованиях входит в состав образовательного процесса (Kimmel, 1996). В идеале, участвуя в экспериментах и изучая исследуемые психологические явления, студенты получают более глубокое понимание исследовательского процесса. Чтобы сделать участие студентов действительно добровольным, кроме непосредственного участия в исследовании им предлагают на выбор несколько вариантов. Однако проблемы все равно существуют. Критики утверждают, что при формировании фондов участников в действительности не соблюдается принцип добровольности, что альтернативные занятия (например, оформление документов) зачастую настолько тягостны, что студентов по сути вынуждают участвовать в исследованиях, а также что приобретаемый опыт скорее скучный, чем поучительный (Kern, 1988). В поддержку такого мнения были проведены исследования. В исследовании Зибера и Сакса (Sieber and Saks, 1989) было обнаружено, что 89% из 366 рассмотренных факультетов содержат фонды, не соответствующие одной или нескольким рекомендациям ассоциации *APA.* Тем не менее многие факультеты стараются сделать так, чтобы студенты могли извлечь пользу из подобных занятий. Например, во время дебрифинга после эксперимента по изучению памяти участникам/студентам могут рассказать о том, как данное исследование связано с главой X из учебника, используемого во вводном курсе. Также на многих факультетах студентам в качестве альтернативы предлагают творческие виды деятельности. Студенты, не участвующие в исследованиях могут: а) наблюдать проводимые эксперименты и записывать свои наблюдения, б) участвовать «за кулисами» в проектах преподавателей факультета (например, кодировать данные), в) участвовать в других мероприятиях научного коллектива или г) слушать лекции ученых, приехавших в данное учебное заведение (Kimmel, 1996). Некоторые исследования показали, что студенты в основном находят участие в экспе-

риментах весьма полезным, особенно если исследователи действительно стараются «привязать» их участие к процессу обучения, протекающему на занятиях по общей психологии (например, Landrum &Chastain, 1999; Leak, 1981).

Ассоциация *APA* (1982, p. 47-48) разработала особые требования к отбору студентов для участия в исследованиях. Вот основные из них:

1) студенты должны быть осведомлены о требованиях курса, прежде чем на него записываться;
2) студенты в первый день занятий по данному предмету должны получить подробное описание требований курса, в том числе ясное описание альтернативных видов деятельности, в случае если они не хотят участвовать в исследованиях;
3) альтернативные виды деятельности должны быть равными участию в исследовании по времени и трудоемкости, а также, аналогично участию в экспериментах, должны иметь образовательное значение;
4) все предложения о проведении исследований с использованием фондов участников должны получать одобрение соответствующего исследовательского совета;
5) особое внимание следует обратить на вежливое обращение со студентами;
6) должна существовать простая и понятная процедура, посредством которой студенты могут пожаловаться на некорректное обращение. Оценка по данному предмету при этом не должна быть снижена;
7) все остальные требования этического кодекса *APA* должны неукоснительно выполняться;
8) на факультете должен быть разработан механизм для периодической проверки правил использования фонда участников.

«Проблема второкурсников» — это только один пример затруднений, связанных с распространением результатов исследования на другие группы. Другой пример касается пола участников. Некоторые из наиболее известных психологических исследований ограничены изучением только мужчин (или, но гораздо реже, только женщин), но распространяют свои выводы на людей обоих полов. Вероятно, самый известный пример — это исследование Лоренса Колберга, посвященное изучению нравственного развития детей. В ходе исследования, проведенного в 1963 г., Колберг просил мальчиков-подростков прочитать, а затем разрешить различные коротко сформулированные нравственные дилеммы. На основании полученных ответов Колберг разработал теорию о шести стадиях морального развития, закрепившуюся в литературе по психологии развития. Человек, находящийся на самой высокой стадии, поступает в соответствии с универсальными принципами справедливости и соблюдения прав личности.

Теория Колберга была подвергнута критике за низкую внешнюю валидность. К примеру, Кэрол Гиллиган утверждала (Gilligan, 1982), что модель Колберга не учитывает половых различий в особенностях мышления и способах принятия моральных решений. Мужчины могут уделять огромное внимание правам личности, а женщины больше внимания обращают на сохранение взаимоотношений между людьми, а следовательно, ответы женщин на некоторые из дилемм Колберга могут не соответствовать столь «высокой ступени развития», как ответы мужчин. По мнению Гиллиган, это связано с недостатками модели Колберга, составившего выборку из подростков исключительно мужского пола.

Также психологи-исследователи обращают внимание на перенос результатов из одной культуры в другую. Например, считается, что в «индивидуалистических» культурах личность ставится выше группы и высоко ценятся личная ответственность и инициатива. С другой стороны, в «коллективистских» культурах группа оценивается выше личности (Triandis, 1995). Выводы исследований, базирующихся на одной культуре, могут не получить универсального применения. Приведем один пример. В США большинство детей учат высоко ценить личные достижения, а, например, в Японии дети считают, что если они выделяются из коллектива, то тем самым снижают ценность остальных членов группы, и поэтому личные достижения ценятся не столь высоко. В одном исследовании было обнаружено, что у американских студентов личные достижения связаны с положительными эмоциями, а у японских — с *отрицательными* (Kitayama, Markus, Matsumoto & Norasakkunkit, 1997). Решить, что положительное отношение к личным достижениям — это общечеловеческая особенность, значит совершить ошибку. Но означает ли это, что во всех психологических исследованиях должны проводиться межкультурные сравнения? Конечно нет. Это лишь означает, что выводы необходимо делать очень внимательно и всегда указывать группу, изучаемую в ходе исследования.

Другие условия

Обладающие внешней валидностью результаты можно распространить не только на другие группы людей, но и на другие условия — они действительны также для других наборов стимулов. Этот вопрос лег в основу критики лабораторных исследований, упомянутой в главе 3: иногда говорят, что такие исследования искусственны и слишком далеки от реальной жизни. Вспомните, как при обсуждении фундаментальных и прикладных исследований рассказывалось, что ответом работающих в лаборатории исследователей на критику явился тезис Аронсона об экспериментальном реализме. Наиболее важно, чтобы люди действительно были вовлечены в исследование, а жизненный реализм при этом уходит на второй план.

Тем не менее попытки изучать психологические явления в реальных жизненных условиях привели к важным открытиям во многих областях психологии. Хороший пример тому — история исследований человеческой памяти. Большую часть XX в. исследования памяти проводились преимущественно в лабораториях, где бесчисленные второкурсники запоминали бесконечные списки слов, бессмысленных слогов, последовательности цифр и т. д. Эти исследования позволили получить большое количество информации об общих механизмах памяти, и это очень важно, но не было ясно, можно ли открытые в лаборатории принципы распространить на ситуации реальной жизни. Переломной точкой стала работа ученого Корнельского университета Ульриха Нейссера «Познание и реальность» (Neisser, *Cognition and Reality,* 1976). В своей работе он указал, что лабораторная традиция в когнитивной психологии, приводя к важным открытиям, тем не менее не способна в достаточной мере пролить свет на процессы обработки информации в реальных условиях, и призвал к исследованиям, обладающим **экологической валидностью,** — исследованиям, важным для каждодневной познавательной деятельности людей, старающихся приспособиться к окружающим условиям. Найссер заявил, что экспериментальная психология «должна приложить серьезные усилия для понима-

ния особенностей познания, протекающего в обычных условиях и в процессе естественной целенаправленной деятельности. Это означает не прекращение лабораторных экспериментов, а переход от исследования переменных, которыми легко управлять, к изучению экологически важных переменных» (p. 7).

На призыв Найссера откликнулись многие (но, конечно, не все) исследователи, занимающиеся вопросами познания. В 1980-1990-х гг. увеличилось число исследований памяти свидетелей (Loftus, 1979) и долговременной памяти — как, например, запоминание содержания школьных курсов, таких как испанский язык (Bahrick, 1984). Сам Найссер провел интересный анализ памяти Джона Дина (Neisser, 1981), главного советника Белого Дома, положившего конец попыткам президента Ричарда Никсона скрыть незаконные действия во время уотергейтского скандала в начале 1970-х гг. Его показания в конгрессе подействовали на развитие скандала и в конечном счете привели к отставке Никсона. Показания Дина заняли 245 страниц текста и были настолько подробными, что некоторые обозреватели назвали его живым диктофоном. Позднее выяснилось, что описанные Дином встречи в Овальном кабинете были записаны также Белым Домом. Сравнение этих записей с показаниями Дина дало Найссеру прекрасную возможность оценить предполагаемую фотографическую память Дина, которая в конце концов оказалась не столь совершенной. Дин смог вспомнить основные вопросы, обсуждавшиеся на таких встречах, но пропустил множество деталей. Для обсуждаемой темы внешней валидности важно, что исследование Найссера отлично показывает, как наши знания о механизмах памяти могут быть расширены с помощью изучения явлений за пределами лаборатории.

Изменение временного контекста

Еще один вопрос, связанный с внешней валидностью, касается долговечности результатов. Одно из самых известных в истории психологических исследований — это исследования конформности, проведенные Соломоном Ашем в 1950-х гг. (например, Ash, 1956). Эксперименты проводились в исторический период, когда в США главенствовали консервативные силы, люди опасались «красной угрозы», а конформность и послушание властям высоко оценивались обществом. Аш обнаружил, что студенты колледжей очень восприимчивы к давлению сверху. Действительно ли это и сегодня, в другом политическом климате? Будет ли фактор воздействия на конформность, найденный Ашем (групповое согласие), главенствовать и в наши дни? Вообще, исследования, посвященные более фундаментальным процессам (например, познанию), выдерживают проверку временем лучше, чем исследования социальных факторов, включенных в определенный исторический контекст.

Обобщая изложенное ранее, можно сказать, что валидность данных увеличивается, если их можно применить к другим людям, другим условиям и другой эпохе. Но должны ли поэтому ученые разрабатывать исследования, включающие множество разных социальных групп, проводить эксперименты в различных условиях и повторять их каждые десять лет? Конечно, нет. Внешняя валидность не определяется по одному исследованию. Она представляет собой еще один пример важности процесса повторения и дополнения. Для ученого, разрабатывающего исследование, важность внешней валидности теряет значение в сравнении с важностью нашей следующей темы.

Внутренняя валидность

Последний вид экспериментальной валидности, описанный Куком и Кэмпбеллом (Cook and Campbell, 1979), называется **внутренняя валидность.** Этот термин обозначает методологическое совершенство эксперимента и отсутствие в нем осложнителей. При обладающем внутренней валидностью исследовании экспериментатор уверен, что результаты, полученные измерением зависимой переменной, непосредственно связаны с независимой переменной, а не с каким-нибудь другим неконтролируемым фактором. В исследовании, отягощенном осложняющими факторами, как мы видели на примере эксперимента с массированной/распределенной нагрузкой, результаты невозможно интерпретировать. Результат может быть связан с влиянием независимой переменной, осложняющей (осложняющих) переменной (переменных) или их комбинации, при этом невозможно выбрать одно из нескольких различных объяснений.

Угроза для внутренней валидности

Внутреннюю валидность исследования может снизить любой неконтролируемый внешний фактор, однако есть несколько вопросов, требующих особого рассмотрения (Cook & Campbell, 1979). Так называемая угроза для внутренней валидности наиболее опасна при отсутствии контрольной группы. Эта проблема иногда возникает в исследованиях по оценке программ (глава 10). Часто такая угроза возникает в длительных исследованиях, в ходе которых проводятся несколько измерений. Например, участники могут пройти предварительное тестирование, подвергнуться особому экспериментальному воздействию, а затем пройти заключительное тестирование. В идеале воздействие должно иметь определенный положительный эффект, который можно определить, обнаружив различия между результатами предварительного и заключительного тестирования. Второй вид угрозы возникает, если сравниваются «неэквивалентные» группы. Эта так называемая проблема отбора испытуемых может переплетаться с другими видами угрозы для внутренней валидности.

Предварительное и заключительное тестирование

Верно ли, что студентам легче изучить общую психологию с помощью компьютеризированного курса, где каждый может продвигаться вперед соразмерно своей собственной скорости? Если в колледже будет проводиться программа по снижению экзаменационной тревожности, как можно показать, что она действует эффективно? Если вы обучите людей различным стратегиям запоминания, улучшит ли это их память? Во всех этих эмпирических вопросах речь идет о том, изменятся ли люди, усвоив некоторый опыт (курс, программа, тренировка памяти). Стандартный способ обнаружить изменения — проверить людей до получения ими данного опыта (провести **предварительное тестирование),** а затем после получения опыта (провести **заключительное тестирование).** Идеальным результатом для только что описанных примеров будет, если люди в ходе заключительного тестирования покажут, что они: а) знают общую психологию лучше, чем вначале, б) меньше волнуются перед экзаменами, чем раньше, и в) демонстрируют улучшение памяти. В ходе

стандартного исследования сравниваются экспериментальная и контрольная группы; последняя при этом не подвергается никакому воздействию:

Экспериментальная группа:	предварительное тестирование	*Воздействие*	заключительное тестирование
Контрольная группа:	предварительное тестирование		заключительное тестирование

При отсутствии подходящей контрольной группы возникает угроза внутренней валидности эксперимента. Предположим, вы хотите оценить эффективность программы помощи студентам колледжа, страдающим от экзаменационной тревоги (т. е. тем, которые имеют неплохие навыки обучения и знают материал, но так сильно волнуются во время экзаменов, что это мешает их успешной сдаче). В начале лекций первокурсники заполняют различные опросники, в том числе один, выполняющий функцию предварительного тестирования для изучения тревожности. Предположим, что оценки могут варьироваться от 20 до 100, и чем больше значение, тем выше уровень тревожности. Студентов, получивших высокие оценки, просят участвовать в программе по снижению экзаменационной тревоги, включающей тренинги на расслабление и выработку учебных навыков, а также другие методики. Спустя три месяца студенты снова выполняют тест на тревожность и его результаты выглядят следующим образом:

Предварительное тестирование	*экспериментальное воздействие*	заключительное тестирование
90		70

Таким образом, средняя предварительная оценка у студентов, выбранных для участия в программе, равна 90, а средняя заключительная — 70. Если предположить, что разница статистически значима, какой вывод можно сделать? Была ли эффективна программа? Чем были вызваны изменения, экспериментальным воздействием, или же могли сказаться другие факторы? Я надеюсь, вы понимаете, что есть несколько способов объяснения таких результатов. Читайте дальше.

История и взросление

Иногда между предварительным и заключительным тестированием вклинивается некоторое событие, не связанное с экспериментальным воздействием, но вызывающее большие изменения. В таком случае угрозой для внутренней валидности исследования является история. Предположим, к примеру, что в колледже из предыдущего примера решили, что система оценок мешает обучению и что знания по всем предметам впредь должны оцениваться по принципу сдал/не сдал. Допустим, что такое решение было принято после проведения предварительного тестирования тревожности в середине экспериментальной программы по ее снижению. Заключительное тестирование может показать значительное снижение тревожности, но очень вероятно, что такой результат будет вызван изменением системы оценок, а не самой программой. Разве вы не станете меньше волноваться по поводу данного курса об исследовательских методах, если оценка не будет иметь никакого значения?

Кроме того, в программе по снижению экзаменационной тревоги студенты-первокурсники принимают участие в самом начале своей учебной деятельности, поэто-

му изменение оценок предварительного и заключительного тестирования может быть связано со **взрослением** студентов по мере привыкания к жизни колледжа. Как вы, вероятно, помните, первый семестр в колледже — это время значительных изменений в жизни студентов. Если исследование занимает продолжительный период, то взросление студентов обязательно скажется на его результатах.

Обратите внимание, что если будет использоваться контрольная группа, то экспериментатор сможет учесть влияние и истории, и взросления. Если будут получены приведенные ниже результаты, то это влияние можно будет исключить, а программу снижения экзаменационной тревожности считать эффективной:

Экспериментальная группа:	предварительное тестирование 90	*экспериментальное воздействие*	заключительное тестирование 70
Контрольная группа:	предварительное тестирование 90		заключительное тестирование 90

С другой стороны, если оценки контрольной группы при заключительном тестировании также снижаются до 70, и история, и взросление, а также оба фактора одновременно могут являться объяснением изменений оценок экспериментальной группы.

Регрессия

Регрессия означает движение в обратную сторону — в данном случае в направлении средней оценки. Поэтому явление, которое я собираюсь описать, иногда называют **регрессией к среднему.** В большом наборе оценок большинство из них группируются вокруг среднего арифметического и лишь некоторые (крайние) оценки будут заметно отстоять от него. Представьте, что вы случайным образом выбираете несколько оценок из нормального распределения, показанного на рис. 5.2. Большинство оценок группируются вокруг среднего, так что если вы выбираете случайным образом, то наиболее вероятно, что вы выберете оценку, близкую к среднему арифметическому (X слева на рис. 5.2). Предположим, однако, что вы выбрали одну, далеко отстоящую от средней оценки (т. е. крайнюю оценку — Y). Верно ли, что при следующем выборе вы, вероятнее всего, получите:

1. Ту же оценку еще раз?
2. Оценку еще более крайнюю, чем первая?
3. Оценку ближе к среднему арифметическому, чем первая?

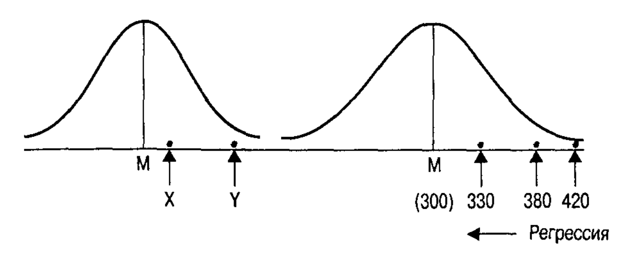


Рис. 5.2. Регрессия к среднему

Я полагаю, вы выберете третий вариант, и это будет означать, что вы понимаете идею регрессии к среднему. Приведем более конкретный пример (относящийся к правой части рис. 5.2). Предположим, вы знаете (на основании нескольких сотен бросков), что в среднем Тэд бросает бейсбольный мяч на 300 футов. Затем он бросает мяч на 380 футов. Если вы будете спорить о его *следующем* броске, на что вы поставите?

1. На 380 футов.
2. На 420 футов.
3. На 330 футов.

И опять я думаю, что вы выберете третий вариант, тем самым еще раз подтвердив, что понимаете идею регрессии к среднему. Но как это связано с нашими предварительными и заключительными исследованиями?

Для многих исследований, включающих предварительное и заключительное тестирование, участники отбираются на основании того, что в ходе предварительной проверки они показали результаты, близкие к крайним. Так, в исследовании экзаменационной тревожности отбирались те участники, чьи показатели тревожности были очень высоки. Заключительное тестирование может показать улучшение показателей, которое возможно будет связано с регрессией, а не с программой улучшения памяти. И снова контрольная группа, состоящая из участников с высокой тревожностью, поможет исследователю заметить наличие регрессии. Например, следующие результаты предполагают, что регрессия могла оказать некоторое влияние[1], но тем не менее воздействие самой программы было более сильным и перекрыло регрессию. Вы понимаете, почему это так?

Экспериментальная группа:	предварительное тестирование 90	*экспериментальное воздействие*	заключительное тестирование 70
Контрольная группа:	предварительное тестирование 90		заключительное тестирование 80

Регрессия может привести к различным проблемам и вероятно именно она была виновна в проведении исследований, ставящих под сомнение эффективность хорошо известной программы *Head Start.* Этот случай будет рассматриваться в главе 10 для иллюстрации проблем, возникающих при осуществлении крупномасштабных федеральных программ.

Тестирование и инструментарий

Тестирование рассматривается как угроза внутренней валидности, когда сам факт проведения предварительного тестирования влияет на послеэкспериментальные оценки. Если проводится многократное тестирование, то может появиться эффект

[1] Обратите внимание на использование выражения «может оказать», а не «оказывает», так как возможно, что оценки контрольной группы изменились с 90 до 80 под влиянием одного из факторов, угрожающих внутренней валидности. Если эти факторы будут исключены, регрессия станет определенной.

привыкания, а кроме того, некоторые аспекты доэкспериментального тестирования могут сделать участников более восприимчивыми к определенным особенностям программы. Например, если объектом эксперимента является компьютеризированный курс по общей психологии, который каждый может изучать с необходимой ему скоростью, то в ходе предварительного тестирования могут проверяться знания по данному предмету и участники могут обратить внимание на вопросы, о которых они ничего не знают, а в дальнейшем в ходе обучения уделять им больше внимания и в результате лучше справиться с заключительной проверкой.

Инструментарий становится проблемой, если между предварительной и заключительной проверкой измерительные инструменты как-либо изменяются. В упомянутом курсе для самостоятельного изучения по общей психологии предварительные и заключительные тесты состоят из разных заданий с одинаковым уровнем сложности. Если, к примеру, заключительный тест будет легче, то его результаты покажут увеличение знаний, не соответствующее действительности. Иногда инструментарий вызывает проблемы, если измерение включает наблюдение. Наблюдатели лучше справятся с работой, попрактиковавшись в ней, а это означает, что инструмент при заключительном тестировании будет отличаться от используемого при предварительном тестировании (в данном случае он будет более точным).

Так же как и проблемы, вызванные историей, взрослением и регрессией, связанные с инструментарием осложнения можно устранить с помощью контрольной группы. Единственным исключением является случай предварительной сенсибилизации, когда экспериментальная группа в ходе заключительного тестирования имеет некоторое преимущество перед контрольной группой, так как полученная при предварительном тестировании информация поможет им в ходе эксперимента сконцентрироваться на пробелах в своих знаниях, тогда как члены контрольной группы не будут иметь такой возможности.

Предварительное оценочное тестирование: план Соломона

Иногда влияние предварительного тестирования можно оценить с помощью особого плана, называемого планом Соломона для четырех групп (Solomon, 1949). Он изображается следующим образом:

Экспериментальная группа 1:	предварительное тестирование	*экспериментальное воздействие*	заключительное тестирование
Экспериментальная группа 2:		*экспериментальное воздействие*	заключительное тестирование
Контрольная группа 1:	предварительное тестирование		заключительное тестирование
Контрольная группа 2:			заключительное тестирование

Сравнение двух экспериментальных групп позволяет определить влияние предварительного тестирования на экспериментальное воздействие, приводящее к изменениям результатов заключительного тестирования. Аналогично сравнение результатов тестирования членов контрольных групп показывает, воздействует ли само предварительное тестирование на выполнение заключительных тестов. Идеальный

результат использования плана Соломона при оценке программы по снижению экзаменационной тревоги выглядит следующим образом:

Экспериментальная группа 1:	предварительное тестирование 90	*экспериментальное воздействие*	заключительное тестирование 70
Экспериментальная группа 2:		*экспериментальное воздействие*	заключительное тестирование 70
Контрольная группа 1:	предварительное тестирование 90		заключительное тестирование 90
Контрольная группа 2:			заключительное тестирование 90

Результаты, свидетельствующие о том, что низкая тревожность при заключительном тестировании вызвана воздействием предварительного тестирования, будут следующими:

Экспериментальная группа 1:	предварительное тестирование 90	*экспериментальное воздействие*	заключительное тестирование 70
Экспериментальная группа 2:		*экспериментальное воздействие*	заключительное тестирование 90
Контрольная группа 1:	предварительное тестирование 90		заключительное тестирование 70
Контрольная группа 2:			заключительное тестирование 90

План Соломона — хороший способ оценки влияния предварительного тестирования, но затраты на его реализацию довольно велики. В сравнении с обычным планом, подразумевающим наличие экспериментальной и контрольной групп, он требует больше участников, а также проведения двух отдельных экспериментальных программ.

Проблемы, связанные с участниками

Угроза внутренней валидности может также исходить от участников исследования. Кук и Кэмпбелл в 1979 г. выделили две связанные с этим проблемы.

Эффект отбора испытуемых

Характерной особенностью экспериментального исследования с управляемой независимой переменной является то, что его участники, находящиеся в разных условиях, эквивалентны по всем показателям, кроме независимой переменной. В следу-

ющей главе вы узнаете, как формировать эквивалентные группы при помощи методов случайного распределения и уравнивания. Если группы не эквивалентны друг другу, может возникнуть **эффект отбора испытуемых.** Предположим, что исследователь на примере двух разделов курса по общей психологии хочет сравнить обычный лекционный курс с другим, состоящим из лекций и дискуссионных занятий. Правила учебного заведения: а) не позволяют исследователю случайным образом разделить студентов на две группы, каждая из которых будет посещать один из курсов, а также б) требуют подробного разъяснения студентам сущности обоих курсов. Таким образом, студенты могут записаться на изучение любого из разделов — в этом и заключается трудность. Допустим студенты, посещающие курс с лекциями и дискуссиями, справились с ним лучше посещающих обычный лекционный курс. Что явилось причиной такой разницы? Связано ли это с организацией курса (наличие дискуссий) или с самими студентами, *выбравшими* этот курс? Может быть, эти студенты имели более развитые речевые способности (а следовательно, были более склонны к обсуждениям), чем другие студенты. Если кратко охарактеризовать сложившуюся ситуацию, то можно сказать, что возникло осложнение, связанное с отбором испытуемых для двух сравниваемых групп.

Эффект выбора может также взаимодействовать с другими факторами, угрожающими внутренней валидности. Например, в исследовании с двумя группами некоторое историческое событие может подействовать на одну группу и не подействовать на другую. В этом случае возникнет осложнитель история χ выбор. Аналогично у двух групп могут быть разные показатели взросления и регрессии, они могут по-разному реагировать на факт тестирования и на них по-разному может влиять инструментарий.

Хорошим примером, демонстрирующим эффект выбора, является (к сожалению) одно из классических исследований по психологии. Это новаторское исследование в области психологии здоровья сегодня известно под названием «язва у "исполнительных" обезьян». Во вставке 5.3 вы найдете описание этого методологически несовершенного (а также сомнительного с этической точки зрения) эксперимента и его последствий.

ВСТАВКА 5.3

Классические исследования- проблема отбора и исполнительные обезьянки

В апрельском номере за 1958 г. журнала «Сайентифик америкэн» была напечатана примечательная статья, озаглавленная «Язва у "исполнительных" обезьян». В ней описывались эксперименты Джозефа В. Брэди из Военного исследовательского института Уолтера Рида. Заинтересованный военными исследованиями, показавшими взаимосвязь между желудочной секрецией солдат во время медицинского осмотра и последующим развитием язвы желудка, Брэди решил исследовать взаимосвязь между эмоциональными и физическими расстройствами путем изучения эмоционального поведения обезьян, принадлежащих к виду макак-резус. К глубокому сожалению ученого, многие из обезьян умерли, а вскрытие показало, что большинство из них страдали язвой желудка. В основном эти обезьяны использовались для

формирования у них условного избегания. В ходе эксперимента их пристегивали к специальным стульям, позволяющим реагировать на идущий к ступням электрический разряд только нажатием рычага.

Брэди предположил, что развитие язвы могло быть связано или с повторяющимся разрядом, или со стрессом, вызванным процедурой выработки условного избегания. Чтобы выявить возможность развития язвы из-за воздействия электрического разряда, применив метод фальсификации, ученый разработал особый эксперимент. Он поместил двух обезьян на соседние стулья; одна из них, «исполнитель» (обратите внимание на аналогию с вечно занятым и испытывающим стресс исполнительным директором), могла избегать разряда, подаваемого каждые 20 секунд, нажав на рычаг в промежуток между двумя разрядами. У контрольной обезьяны рычаг не работал, и она испытывала разряд каждый раз, когда управляющая обезьяна пропускала 20 секунд и получала разряд. Таким образом, обе обезьяны испытывали разряд одинаково часто, но только одна из них могла контролировать ситуацию. Так продолжалось в течение 6 часов, а затем следовал 6-часовой отдых. Далее, через 24 часа эксперимент возобновлялся - проводились две попытки по 6 часов с двумя периодами 6-часового отдыха (при этом обезьяны ни разу не покидали стульев).

Через двадцать три дня (!) обезьяна-исполнитель умерла, после чего обнаружили, что у нее была язва желудка. У контрольной обезьяны язва не была обнаружена. Брэди повторил эксперимент с другой парой обезьян и получил те же результаты. В дальнейшем он привел отчет об исследовании с четырьмя парами животных (Brady, Porter, Conrad & Mason, 1958), в котором сделал вывод, что психологический стресс, вызванный ситуацией, когда приходится управлять не только собственной судьбой, но также и судьбой подчиненного, может привести к проблемам со здоровьем, в данном случае к язве желудка. Догадываетесь, почему армия профинансировала это исследование?

Описание данного исследования было приведено во многих учебниках по общей психологии, а публикация в «Сайентифик америкэн» привлекло к нему еще большее внимание. Однако тщательное изучение использованного Брэди метода показало, что в нем имело место осложнение из-за процедуры выбора субъектов. Брэди разделил обезьян на две группы не случайным образом, а на основании того, как быстро они освоили процедуру избегания. Сначала все восемь обезьян испытывались как исполнители, а затем те, которые реагировали быстрее, были сделаны исполнителями на время эксперимента. Брэди не знал, что животные различаются по присущей им эмоциональности и более эмоциональные быстрее реагируют на болевое воздействие, поэтому он невольно сделал исполнителями более эмоциональных животных (т. е. предрасположенных к язве желудка), а более спокойных поместил в контрольные условия.

Первым на эту проблему указал Вейсс в 1968 г. Его исследования с крысами были лучше организованы (например, Weiss, 1977) и показали результаты, *противоположные* результатам Брэди. Как вы узнаете из рассказа об эквивалентных контрольных группах в главе 7, Вейсс обнаружил, что язва развилась у меньшего количества животных, имеющих возможность контролировать электрический разряд, и у большего количества животных, не имеющих такой возможности.

Еще один аспект исследования Брэди, который вы, вероятно, уже отметили, заключается в том, что сегодня Комиссия по использованию животных потребует существенных изменений в процедуре исследования. Изучение взаимосвязи между стрессом и физическими расстройствами является жизненно важным, но его можно исследовать с помощью менее жестокой процедуры, когда обезьян не будут привязывать к стульям и оставлять практически обездвиженными в течение нескольких недель.

Истощение

Испытуемые не всегда участвуют в экспериментах вплоть до их завершения. Некоторые исследования могут проводиться довольно долгое время, и люди могут прекратить свое участие, потерять интерес или даже умереть. В ходе некоторых исследований участники могут испытывать дискомфорт и воспользоваться правом отказа от участия в дальнейших экспериментах. Поэтому вследствие различных причин в начале исследования в нем могут участвовать 100 человек, а в конце остаться только 60. Эту проблему иногда называют «смертностью» испытуемых, но я предпочитаю не использовать столь печальное выражение и вместо него применяю термин **истощение.** Истощение является проблемой, так как если одни люди более склонны к тому, чтобы бросить исследование, чем другие, то в конце исследования состав участников качественно будет отличаться от исходного. Данная проблема сходна с проблемой выбора, поскольку в результате группа, начинающая исследование, становится неэквивалентной группе, завершающей его. Чтобы проверить, различаются ли продолжающие исследование учестники и покидающие его, можно сравнить оценки предварительного тестирования или другие показатели, полученные до начала экспериментов, у обеих групп. Если покинувшие исследование и продолжающие участвовать в нем испытуемые не различались в начале, то общие выводы в конце исследования не пострадают, даже несмотря на потери в связи с истощением.

На этом мы завершаем введение в экспериментальный метод. В трех следующих главах будут подробно рассматриваться вопросы, связанные с этой темой. Глава 6 откроется описанием различий межсубъектых и внутрисубъектных планов (или планов с повторяющимися измерениями). Далее в ней будут обсуждаться проблемы, связанные с управлением экспериментальными исследованиями, в том числе вопросы создания эквивалентных групп для экспериментов с межсубъектными планами, контроля за эффектом последовательности для внутрисубъектных планов, а также искажения, возникающие из-за того, что и экспериментаторы и участники — это живые люди. Главы 7 и 8 посвящены рассмотрению различных планов исследований, начиная от планов с одной независимой переменной (глава 7) и заканчивая планами с множеством независимых переменных, известными как факторные планы (глава 8).

Резюме

Основные черты экспериментальных исследований

Психологический эксперимент включает в себя задание независимых переменных, контроль за внешними переменными и измерение зависимых переменных. Независимые переменные возникают при проведении сравнений и при создании экспериментальных условий, находящихся под непосредственным контролем исследователя. Управляемые независимые переменные подразумевают помещение участников в разные условия, предложение им разных заданий или разных инструкций для выполнения заданий. Внешние переменные — это факторы, которые сами по себе не интересуют исследователя. Невозможность управлять ими приводит к проблеме, называемой осложнением. При наличии осложнителя результаты могут быть

вызваны как независимой переменной, так и осложнителем. Зависимые переменные — это характеристики поведения, измеряемые в ходе исследования. Они должны быть точно (операционально) определены.

Управляемые и субъективные переменные

В некоторых психологических исследованиях сравниваются группы участников, определенным образом отличающихся друг от друга еще до начала эксперимента (например, могут различаться пол, возраст, интровертированность). В таком случае интересующая исследователя независимая переменная выбирается для эксперимента, а не регулируется непосредственно и называется субъективной переменной. Психологические исследования часто проводятся с использованием и управляемых, и субъектных переменных. В исследованиях, достаточно хорошо контролируемых экспериментатором, выводы о причинах и следствиях могут быть сделаны только при использовании управляемых переменных, в исследованиях с субъективными переменными это невозможно.

Валидность экспериментальных исследований

Есть четыре стороны валидности психологических исследований. В валидном исследовании правильно применяется статистический анализ (валидность статистических выводов), адекватно определяются независимые и зависимые переменные (конструктная валидность), его результаты можно распространить за пределы проведенного эксперимента (внешняя валидность), а кроме того, оно не содержит осложняющих переменных (внутренняя валидность).

Угроза для внутренней валидности

Угрозу для внутренней валидности эксперимента могут представлять различные факторы. В плохо контролируемых исследованиях, включающих сравнение результатов предварительного и заключительного тестирования, часто появляются такие осложняющие факторы, как история, взросление, регрессия, тестирование и инструментарий. Проблема отбора возникает, когда сравниваются две неэквивалентные группы участников, различия между которыми существуют еще до начала исследования (например, как в исследовании Брэди, посвященном изучению образования язвы желудка у «исполнительных» обезьян). Так же проблема отбора может взаимодействовать с другими факторами, угрожающими внутренней валидности. В долгосрочных исследованиях истощение может привести к возникновению особой разновидности проблемы отбора — небольшая группа участников, оставшихся к концу исследования, может заметно отличаться от более многочисленной группы, набранной в начале.

Задания для повторения

Выбор ответа

1. В главе 4 в качестве примера использования номинальной шкалы было описано исследование Шехтера, показавшее, что ожидающие воздействия сильного электрического шока люди предпочитают ждать вместе с другими. Ис-

следование можно подвергнуть критике за то, что все его участники были женщинами. Это будет критикой:

1) конструктной валидности исследования; .
2) внешней валидности;
3) внутренней валидности;
4) экологической валидности.

2. Для участия в программе по снижению повышенного кровяного давления отобрали 50 добровольцев, которым остался один год до выхода на пенсию. Через год после выхода людей на пенсию все 50 человек продолжали участвовать в программе, а их кровяное давление снизилось. Какой угрожающий внутренней валидности фактор затрудняет интерпретацию данных этого исследования?

1) истощение;
2) отбор испытуемых;
3) инструментарий;
4) история.

3. В исследовании памяти экспериментатор хочет показать, что использование зрительных образов улучшает запоминание. Одних участников попросили придумывать образы, а других — просто повторять каждое слово по два раза. Поскольку формирование образов занимает больше времени, чем повторение слов, создающей образы группе слова демонстрировали с более низкой скоростью (4 секунды на каждое), чем группе, повторяющей слова (2 секунды на слово). Какое из следующих утверждений верно?

1) скорость показа осложняется инструктивной независимой переменной;
2) скорость показа является независимой переменной;
3) исследование обладает внутренней валидностью, но не обладает внешней;
4) исследование хорошо спланировано. Инструктивная переменная является управляемой незавивимой переменной.

4. Какое из следующих утверждений об известном исследовании с куклами бобо верно?

1) все независимые переменные были управляемыми;
2) исследователи управляли фактором «бобо», используя одну и ту же надувную куклу как для взрослых, так и для детей;
3) дети имитировали модели, только если они были представлены в виде мультфильмов;
4) управляемая независимая переменная являлась ситуативной переменной.

5. В исследовании программы снижения тревоги средняя предварительная оценка в экспериментальной группе была 79 (максимальная = 100), а средняя заключительная — значительно более низкой: 68. Наиболее вероятно, что такое изменение было вызвано:

1) регрессией;

2) взрослением;
3) эффективностью программы;
4) любой из вышеперечисленных факторов (или их комбинация) мог вызвать подобное изменение.

Короткие эссе

1. На примере исследования тревожности проиллюстрируйте разницу между а) управляемыми и б) субъективными независимыми переменными.
2. На различных примерах покажите разницу между ситуативными, рабочими и инструктивными независимыми переменными.
3. Что такое осложнитель и почему его наличие затрудняет интерпретацию результатов исследования?
4. Считается, что в исследовании с использованием субъективных переменных выводы о причинах явлений не могут быть сделаны. Почему?
5. Опишите условия, при которых может снизиться валидность статистических выводов эксперимента.
6. Опишите три вида условий, при которых может снизиться внешняя валидность.
7. Охарактеризуйте различия внутренней и внешней валидности исследования.
8. Объясните, каким образом наличие контрольной группы может снизить угрозу внутренней валидности. В качестве примера используйте историю, взросление или регрессию.
9. Объясните, как можно использовать план Соломона дня оценки влияния предварительного тестирования участников.
10. На примере исследования Брэди проиллюстрируйте проблему отбора.

Упражнения

Упражнение 5.1. Выявление переменных

В каждом из перечисленных ниже исследований выявите независимую (независимые) переменную (переменнее), значения независимой (независимых) переменной (переменных) и зависимую (зависимые) переменную (переменные). Определите, являются ли независимые переменные управляемыми или неуправляемыми. Укажите, какая шкала использовалась для измерения зависимых переменных.

1. В исследовании когнитивных карт сравнивалась способность первокурсников и студентов старших курсов точно указывать местоположение зданий кампуса. Некоторые здания располагаются в центре кампуса поблизости от дорог, по которым часто ходят студенты, а другие — на периферии студенческого городка. Участников исследования попросили оценить (по шкале от 1 до 10) степень уверенности в том, что они правильно указали направление, и записать количество ошибок (в процентах).

2. В исследовании эффективности нового лекарства для лечения депрессии одни участники принимали его, а другие думали, что принимают. Третья группа участников не подвергалась экспериментальному воздействию. По окончании программы участники заполнили Опросник депрессии Бека, а затем опытные наблюдатели оценили (по 10-балльной шкале) уровень их депрессии.
3. В исследовании обусловливания, проведенном И. П. Павловым, у голодных животных вырабатывался условный рефлекс слюноотделения при ЗВУКОВОМ сигнале, связанном с получением пищи. Одним животным, прежде чем выдать еду, предъявляли сигнал, а затем снова выключали его. На других животных воздействовали сигналом до самого кормления. Третья группа животных получала еду до включения сигнала. Экспериментаторы отмечали начало саливации и ее объем для определенного промежутка времени.
4. В исследовании по психолингвистике развития двух-, трех- и четырехлетним детям показывали кукол и просили разыгрывать отдельные сценки, чтобы определить, способны ли они использовать определенные грамматические правила. В одних случаях детей просили разыгрывать сценки, используя активный залог (Эрни побил Берта), а в других — используя пассивный залог (Эрни был побит Бертом). Оценивалось, правильно ли дети разыгрывают сценки (две оценки) и как быстро они приступают к игре.
5. В исследовании способностей к запоминанию лабиринта одни крысы на протяжении 30 попыток получали корм после прохождения лабиринта, другие корма не получали, третьих не кормили на протяжении первых 15 попыток, но кормили в конце следующих 15 раз, а последнюю группу крыс кормили первые 15 раз и не кормили вторые 15 попыток. Исследователь отмечал все сделанные ошибки (неправильные повороты) и время прохождения лабиринта.
6. В исследовании, посвященном готовности к оказанию помощи, к посетителям торгового центра подходит либо хорошо одетый студент, либо неряшливый и спрашивает, где находится туалет или ближайший магазин *Kmart*. Расположившийся неподалеку экспериментатор отмечает, была ли оказана помощь.

Упражнение 5.2. Обнаружение осложнителя (осложнителей)

Для каждого из приведенных ниже исследований определите независимые и зависимые переменные, значения, принимаемые независимыми переменными, а также укажите хотя бы одну внешнюю переменную, контроль за которой не был адекватным (т. е которая создала осложнитель). Для выполнения задания воспользуйтесь подходом, показанным в табл. 5.2.

1. Фирма, проводящая тестирование, пытается выяснить, можно ли новой клюшкой для гольфа (клюшка 1) посылать мяч дальше, чем это делают с помощью клюшек трех других конкурирующих торговых марок (клюшки 2-4). Для этого набрали 20 добровольцев — мужчин, профессионально играющих в гольф. Каждый гольфер наносит 50 ударов клюшкой 1, затем 50 клюшкой 2, далее 50 клюшкой 3 и 50 клюшкой 4. Для большего реализма эксперимент

проходит на настоящей площадке для гольфа. Используются первые четыре лунки — участник совершает первые 50 ударов от первой метки для мяча, затем 50 от второй и т. д. Первые четыре лунки находятся на расстоянии 380-400 ярдов, попадание в каждую дает 4 очка.

2. Исследователь хочет определить способность страдающих шизофренией пациентов определять длительность различных событий. Выдвинута гипотеза, что сильный шум будет снижать точность. Проверка участников проводится двумя способами: одних участников исследуют в условиях «тишины» в маленькой звуконепроницаемой комнате, которую используют для проверки слуха, а других — в комнате медицинского персонала, где играет музыка с постоянной (и довольно высокой) громкостью. По причине существующего расписания пациенты, которым не разрешено самостоятельно покидать палаты (т. е. немного более опасные), могут участвовать в исследовании *только* в понедельник, а те, которые могут свободно передвигаться (т. е. немного менее опасные), — *только* в четверг. Кроме того, в четверг проводится проверка слуха, поэтому звуконепроницаемая комната может использоваться только в понедельник.
3. Экспериментатор хочет определить, улучшится ли память у людей, использующих визуальные образы. Участники (женщины) разделяются на две группы — в первой используют технику создания зрительных образов, а во второй запоминают путем механического повторения. Первой группе предлагают для изучения список из 20 существительных, обозначающих конкретные понятия (для них легче, чем для абстрактных, сформировать образы), а второй — 20 абстрактных слов (таких, которые особенно легко произносить, а следовательно, повторение не составит труда), имеющих такую же частоту общего использования, как и конкретные слова. Для согласования метода представления слов со способом их запоминания участникам из первой группы слова демонстрируют визуально (на экране компьютера). Для контроля возможной «компьютерофобии» участники из второй группы также сидят у экранов компьютеров, запрограммированных на чтение слов. После прослушивания соответствующих наборов слов участникам дается 60 секунд на их воспроизведение в произвольном порядке.
4. Социальный психолог изучает готовность к оказанию помощи другим людям. Он познакомился с двумя аспирантами, готовыми участвовать в проведении экспериментов. Первый (Нед) хорошо одевается, а второй (Тед) не слишком заботится о своей внешности. В ходе эксперимента к посетителям торгового центра подходит либо хорошо одетый Нед, либо потрепанный Тед и просит двадцатипятицентовик на кофе. Находящийся поблизости экспериментатор записывает, дают ли покупатели деньги. Исследование проводится с 8 до 9 часов вечера, причем Нед работает в понедельник, а Тед — в четверг.

Упражнение 5.3. Операциональные определения

В главе 3 вы познакомились с операциональными определениями и выполнили упражнение на формулировку таких определений для нескольких конструктов, используемых в психологических исследованиях. В данном упражнении вы буде-

те играть роль экспериментатора, разрабатывающего исследование. Для каждой из четырех приведенных ниже гипотез:

а) определите независимую (независимые) переменную (переменные), решите, сколько значений независимой (независимых) переменной (переменных) вы будете использовать, и задайте эти значения;

б) определите зависимую переменную для каждого исследования и

в) дайте операциональные определения независимым и зависимым переменным.

1. Люди скорее помогут нуждающемуся, если необходимость помощи не будет вызывать сомнений.

2. Способность концентрироваться на задании снижается, если человек чувствуют себя в тесноте из-за окружения других людей.

3. Хорошие игроки в боулинг лучше играют в присутствии зрителей, а посредственные — хуже.

4. Животные быстрее запоминают сложный лабиринт, если они испытывают умеренное возбуждение. Они справляются с заданием хуже при сильном или слабом возбуждении. Если лабиринт легкий, способность к его прохождению постепенно повышается при нарастании возбуждения.

ГЛАВА 6

Проблемы контроля при экспериментальных исследованиях

Обзор задач главы

В главе 5 вы познакомились с важнейшими особенностями экспериментального метода — управлением независимой переменной, контролем за всеми остальными факторами и измерением зависимой переменной. В этой главе мы начнем изучать два основных вида экспериментальных планов: межсубъектный план, при котором разные группы участников становятся источниками разных значений независимой переменной, и внутрисубъектный план, в котором на группе одних и те же участников собираются данные всех значений независимой переменной. Вы узнаете, что каждый план имеет свои преимущества и недостатки, а также поймете, каким образом связанные с видом плана проблемы (проблема эквивалентных групп для межсубъектного плана и проблема последовательности для внутрисубъектного плана) должны тщательно контролироваться. В заключительной части главы ставится вопрос искажений и способов контроля за ними. Изучив данную главу, вы:

- сможете отличать межсубъектный план от внутрисубъектного;
- научитесь понимать, каким образом метод случайного распределения решает проблему эквивалентных групп в межсубъектных планах;
- узнаете, в каких случаях для создания эквивалентных групп нужно использовать уравнивание, а не случайное распределение;
- сумеете различать эффект прогрессии и эффект передачи во внутрисубъектных планах и поймете, почему позиционное уравнивание лучше решает первую проблему, чем последнюю;
- сможете описать различные варианты позиционного уравнивания, применяемые в зависимости от того, если участники исследуются или один, или более раз при каждом наборе условий;
- сможете рассказать о видах меж- и внутрисубъектных планов, используемых в психологии развития, и охарактеризовать связанные с их использованием проблемы;
- сможете объяснить причины появления искажения, вызванного экспериментатором, и рассказать о способах контроля за ним;
- сможете описать причины возникновения искажения, вызванного участниками, и способы контроля за ним.

В своем выдающемся учебнике по экспериментальной психологии Р. С. Вудвортс сразу после знакомства с известным сегодня разделением переменных на независимые и зависимые указывает на особую важность контроля за остальными факторами для экспериментальных исследований. По его словам, «используется ли одна или несколько независимых переменных, очень важно поддерживать все остальные условия постоянными. В противном случае вы не сможете связать наблюдаемое явление ни с одной из возможных причин. Психолог должен быть готов к трудностям, связанным с выполнением данного требования...» (Woodworth, 1938, р. 3). С некоторыми из этих проблем мы уже познакомились. Обсуждаемые в предыдущей главе общий вопрос осложнений и конкретные случаи угрозы внутренней валидности в основном связаны с проблемой контроля за внешними факторами. В данной главе мы рассмотрим другие аспекты контроля: проблему создания эквивалентных групп в исследованиях с несколькими группами испытуемых, эффект последовательности в экспериментах, где участники исследуются более одного раза, и проблемы искажений.

Вспомните, что независимая переменная должна принимать по крайней мере два значения, так чтобы исследователь мог сравнивать условие *А* с условием *В*. Испытуемые могут исследоваться при значении *А, В* или при обоих значениях. Если они исследуются при значении *А или* при значении *В,* но не при обоих, то используется **межсубъектный план.** Это название было дано, так как сравнение значений *АиВ* определяет различие *между* двумя группами испытуемых. Если же каждый участник исследуется при значениях *А и В,* то можно сказать, что оба значения имеют место *«внутри»* одного испытуемого, поэтому такой план называется **внутрисубъектным планом** (или иногда планом с повторяемыми измерениями).

Межсубъектные планы

Иногда межсубъектные планы используются потому, что это просто необходимо. Если, например, независимая переменная является субъектной, то другого выбора нет. Исследование, в котором сравниваются интроверты и экстраверты, предполагает участие двух групп испытуемых. Пока ученые не могут собрать несколько личностей в одну, у которой бы одна сторона была интровертированной, а другая — экстравертированной, есть только одна возможность — сравнивать различные группы людей. Один из немногих случаев, когда субъектная переменная не должна быть межсубъектной, — сравнение особенностей поведения, проявляющихся в различном возрасте. При этом одни и те же люди изучаются в различные периоды их жизни. Другой случай — когда в качестве субъектной переменной рассматривается семейное положение. В такой ситуации люди исследуются до и после замужества или развода. Однако в большинстве случаев субъектная переменная требует использования межсубъектного плана.

Использование межсубъектного плана необходимо в некоторых исследованиях с определенными управляемыми независимыми переменными. Иногда получается, что испытуемые, принимая участие в исследовании с одним значением независимой переменной, приобретают опыт, не позволяющий им продолжать эксперименты при других значениях переменной. Это часто происходит в исследованиях

по социальной психологии, а также в исследованиях с использованием мистификации. Рассмотрим эксперимент по изучению воздействия физической привлекательности подсудимого на определяемый ему срок наказания, проведенный Сигаллом и Островым в 1975 г. Студентам колледжа раздали описание преступления и попросили определить срок тюремного заключения для совершившей его женщины. Использовались две независимые управляемые межсубъектные переменные. Первая — вид преступления: либо ограбление, при котором «Барбара Хелм» забралась в соседскую квартиру и украла 2200 долларов, либо мошенничество, при котором «Барбара» «вошла в доверие к холостяку средних лет и убедила его вложить 2200 долларов в несуществующую фирму» (Sigall & Ostrove, 1975, р. 412). Второй переменной была физическая привлекательность «Барбары». Одни участники видели фотографию очень привлекательной «Барбары», другие — непривлекательной (для фотографий позировала одна и та же женщина), а контрольная группа вообще не видела фотографий. Был получен интересный результат: если преступление было ограблением, то привлекательность играла на руку подсудимой: привлекательная «Барбара» в среднем получила более *легкое* наказание (2,8 лет); чем непривлекательная (5,2 года) или контрольная (5,1 года). Но результат был абсолютно противоположным, если преступление было мошенничеством: очевидно, думая, что «Барбара» использовала свою привлекательность для совершения преступления, участники вынесли привлекательной «Барбаре» более строгое наказание (5,5 лет), чем непривлекательной (4,4) и контрольной (4,4).

Из описания исследования понятно, почему его необходимо проводить с использованием межсубъектных переменных. Опыт, полученный испытуемыми в условиях, когда привлекательная «Барбара» совершает воровство, несомненно повлияет на них и не позволит им «начать с нуля» эксперимент, в котором непривлекательная «Барбара» совершает мошенничество. В некоторых исследованиях участие в эксперименте с одним условием делает невозможным участие того же человека в эксперименте со вторым условием. Иногда необходимо, чтобы при каждом условии изучались новые испытуемые.

Преимущество межсубъектного плана заключается в том, что каждый участник начинает эксперимент «с нуля» и не имеет никакой информации о процедуре исследования, а главный недостаток — в том, что необходимо набрать большое количество испытуемых, всех исследовать и для всех провести дебрифинг. Следовательно, этот вид плана требует приложения больших усилий со стороны исследователей. Моя докторская диссертация, посвященная изучению памяти, включала пять различных экспериментов, требующих использования межсубъектных планов, и через мою лабораторию прошло более 600 студентов, прежде чем проект был завершен!

Другой недостаток межсубъектных планов состоит в том, что разница в поведении при различных экспериментальных условиях может быть вызвана независимыми переменными, а может быть связана с различиями между двумя группами. Чтобы устранить возможные осложнители, необходимо создать так называемые **эквивалентные группы** — группы, равные друг другу во всем, кроме значения независимой переменной. Количество эквивалентных групп в межсубъектном исследовании соответствует количеству различных условий этого исследования, и каждая из групп изучается при каждом из условий.

Проблема создания эквивалентных групп

Есть два основных способа создания эквивалентных групп для проведения межсубъектных экспериментов. В идеальном случае используется случайное распределение, второй способ — уравнивание.

Случайное распределение

Во-первых, необходимо понимать, что случайный отбор и случайное распределение - это не одно и то же. Случайный отбор, описанный в главе 4, — это процедура, направленная на отбор добровольцев для участия в исследовании. Случайное распределение представляет собой метод разделения отобранных участников на группы. При **случайном распределении** каждый доброволец имеет равные шансы попасть в каждую из групп.

Задача случайного распределения — равномерно распределить по группам факторы индивидуальных различий, способные исказить исследование. Предположим, вы сравниваете две скорости показа слов, используемых при исследовании памяти. Далее допустим, что склонные к тревожности участники не столь хорошо справились с заданием, как спокойные, но вам этот факт неизвестен. Одним участникам каждое слово показывают в течение 2-х секунд, а другим — 4-х. Делается предположение, что группа, просматривающая слова по 4 секунды каждое, запомнит их лучше. Ниже представлены гипотетические данные, которые могут быть получены в ходе подобного исследования. Каждое число означает количество запомненных слов из 30, включенных в список. После номера участника я поместил в скобках буквы «Т» или «С», чтобы показать, тревожный или спокойный данный участник. Показатели тревожных участников выделены цветом.

Участник	Показ по 2 с	Участник	Показ по 4 с
S1(C)	16	S9(C)	23
S2(C)	15	S10(C)	19
S3(C)	16	S11(C)	19
S4(C)	18	S12(C)	20
S5(C)	20	S13(C)	25
S6(T)	10	S14(T)	16
S7(T)	12	S15(T)	14
S8(T)	13	S16(T)	16
CA	15,00	CA	19,00
CO	3,25	CO	3,70

Внимательно изучив данные, вы заметите, что три тревожных участника в каждой группе хуже справились с заданием, чем пять спокойных, но поскольку каждая группа содержит равное число тревожных испытуемых, снижение количества запомненных слов, вызванное тревожностью, во обеих группах примерно одинаково. Таким образом, интересующее нас различие между двумя скоростями показа слов (показ слов по 4 с каждое приводит к лучшему запоминанию — 19,15) сохраняется.

Случайное распределение не гарантирует, что в каждой группе будет равное число тревожных участников, но в целом процедура приводит к равномерному распределению потенциальных осложнителей по разным группам. Это особенно характерно для групп, состоящих из большого количества людей. По сути, чем больше количество испытуемых, тем больше вероятность, что методом случайного распределения будут созданы эквивалентные группы. Если группы эквивалентны, а все остальное находится под контролем, то вы находитесь в том завидном положении, когда можно утверждать, что различия между группами вызваны независимой переменной.

Вы, наверное, думаете, что процесс случайного распределения весьма прост и для распределения участников по группам нужно лишь использовать таблицу случайных чисел или, в исследованиях с двумя группами, бросать монетку. К сожалению, в результате такой процедуры ваши группы почти наверняка будут различаться по количеству. Представьте себе такое печальное развитие событий: вы проводите исследование с 20 испытуемыми, разделенными на две группы по 10 человек в каждой, и решаете бросать монетку: если орел, то участник попадает в группу *А,* а если решка, то в группу *В.* Но что если решка выпадет все 20 раз?

Чтобы произвести случайное распределение участников по группам и получить одинаковое число людей в каждой группе, можно использовать **блоковую рандомизацию** — процедуру, гарантирующую, что каждому условию исследования случайным образом приводится в соответствие участник, прежде чем какое-либо из условий встречается второй раз. Каждый «блок» содержит все условия исследования, представленные в случайном порядке. В табл. 6.1 показано, как это можно сделать вручную, хотя на практике исследователи часто используют простую компьютерную программу, которая следит за тем, чтобы последовательность условий удовлетворяла требованиям блоковой рандомизации.

Таблица 6.1

Блоковая рандомизация

Блоковая рандомизация, используемая для создания эквивалентных групп, создает блоки, каждый из которых содержит все условия эксперимента. В пределах одного блока условия распределены случайным образом. Ниже представлена блоковая рандомизация, проведенная для исследования, в котором сравнивается влияние четырех скоростей показа слов на их запоминание.

Шаг1. Определите, сколько людей вы будете обследовать. Если вы хотите изучать равное количество людей при каждом из условий, общее количество получается умножением этого значения на количество условий (в данном случае четыре). Предположим, вам нужно 80 человек, но 20 для каждой скорости показа.

Шаг 2. Пронумеруйте четыре условия от 1 до 4. Каждый блок будет содержать случайную последовательность этих чисел.

Шаг 3. Обратитесь к таблице случайных чисел, просмотрите ее по рядам или столбцам и выберите числа от 1 до 4. Выберите каждое число по одному разу, прежде чем делать это повторно. Предположим к примеру, что участок таблицы случайных чисел имеет следующий вид. Вы начинаете выбор со второго ряда и просматриваете таблицу слева направо:

Окончание табл. 6.1

	2	2	1	7	6	8	6	5	8	4	6	8	9	5
	$\underline{1}$	9	$\underline{3}$	6	1	7	5	9	$\underline{4}$	6	$\underline{1}$	$\underline{3}$	7	9
→	1	6	7	7	$\underline{2}$	$\underline{3}$	0	$\underline{2}$	7	7	0	9	6	$\underline{1}$
	7	8	0	$\underline{3}$	7	6	7	$\underline{1}$	6	1	$\underline{2}$	0	$\underline{4}$	4
	0	$\underline{3}$	$\underline{2}$	8	$\underline{1}$	$\underline{2}$	2	6	0	8	7	$\underline{3}$	3	7

Я подчеркнул последовательности случайных чисел, которые необходимо выбрать в данном случае. Таким образом, первый блок будет 1-3-4-2. 1, 3 и 4 заданы в таблице, а если выбраны эти три числа, то четвертым должно быть число 2, поэтому нет необходимости искать его в таблице. Второй блок будет 1-3-2-4 и т. д. Чтобы охватить 80 участников, вам потребуется выбрать 20 блоков по 4 числа в каждом.

Шаг 4. Запишите все последовательности в таблицу.* По окончании обследования каждого из участников вычеркивайте один из 80 номеров таблицы. После того как вы преобразуете числа из таблицы случайных чисел в числа, обозначающие четыре условия эксперимента, и обследуете шесть участников, участок построенной вами таблицы будет иметь следующий вид:

Блок 1 ~~1 3 4 2~~

Блок 2 ~~1 3~~ 2 4

Блок 3 3 2 1 4

• •

• •

• •

Блок 20 2 4 3 1

* Никогда не обследуйте участников, пока не создана таблица, в которой точно указано, как каждый из них должен обследоваться.

Уравнивание

Если в эксперименте принимает участие небольшое количество испытуемых, случайное распределение не всегда помогает создать эквивалентные группы. Приведем пример, который покажет, как это может быть. Возьмем рассмотренное ранее исследование влияния скорости показа слов на их запоминание и допустим, что изученные вами данные получены при использовании случайного распределения, т. е. что в каждую группу попало по пять спокойных и три тревожных испытуемых. Однако *возможно,* что в результате случайного распределения все шесть тревожных участников окажутся в *одной* группе. Это маловероятно, но все же может произойти (так же как возможно, что обычная монета упадет решкой вверх 10 раз подряд). В таком случае результаты могут быть следующими[1]:

[1] Такие же результаты могут быть получены, если экспериментатор не использует случайное распределение, а просто составляет из первых восьми человек группу, в которой слова демонстрируют по 2 секунды, а из вторых восьми группу, в которой их показывают по 4 секунды. Возможно, что более тревожные студенты не сразу захотят участвовать в исследовании, увеличив тем самым вероятность того, что они попадут во вторую группу.

Участник	Показ по 2 с	Участник	Показ по 4 с
S1(C)	15	S9(C)	23
S2(C)	17	S10(C)	20
S3(C)	16	S11(C)	16
S4(C)	18	S12(C)	14
S5(C)	20	S13(C)	16
S6(T)	17	S14(T)	16
S7(T)	18	S15(T)	14
S8(T)	15	S16(T)	17
CA	17,00	CA	17,00
CO	1,69	CO	3,07

Такие результаты, конечно, существенно отличаются от первого примера. Исследователь не сможет отвергнуть нулевую гипотезу (17=17) и заключить, что запоминание улучшается при более низкой скорости показа слов (как в предыдущем примере). Он будет весьма удивлен происходящим, ведь участники были отобраны случайным образом, а его прогноз о том, что запоминание улучшится при более низкой скорости показа, определенно имеет смысл. Так что же было неправильно?

Произошло то, что случайное распределение «случайно» привело к созданию двух неэквивалентных групп: одной, состоящей исключительно из спокойных людей, и второй, членами которой стали преимущественно тревожные люди. Возможно, при показе слов по 4 секунды действительно улучшается запоминание, но в данном исследовании разница стерлась, так как среднее арифметическое для группы, в которой скорость показа составила 2 секунды, выросло, поскольку оценки спокойных участников были относительно высоки, а его значение для группы, где скорость показа была 4 секунды, уменьшилось под влиянием тревожности. Другими словами, неудавшаяся попытка создать эквивалентные группы с помощью случайного распределения привела к ошибке 2-го типа (в действительности скорость показа влияет на запоминание, но в данном исследовании не удалось это обнаружить). Еще раз отметим, что вероятность создания эквивалентных групп с помощью случайного распределения возрастает с увеличением размера выборки.

Решить проблему эквивалентных групп в подобной ситуации можно с помощью процедуры уравнивания. При **уравнивании** испытуемые группируются по принципу обладания какими-либо особенностями (например, в зависимости от уровня тревожности), а затем случайным образом распределяются по разным экспериментальным группам. В рассмотренном выше исследовании памяти «уровень тревожности» можно назвать **переменной уравнивания.** У участников эксперимента с помощью надежного и валидного теста определят тревожность, имеющих сходные оценки людей объединят по парам, а затем одного из них случайным образом определят в группу со скоростью показа 2 секунды, а другого — в группу со скоростью показа 4 секунды. В табл. 6.2 показано, как проводить уравнивание для эксперимента с двумя группами.

Таблица 6.2

Использование процедуры уравнивания

Руководитель исследования с двумя группами, посвященного способности к решению задач, хочет узнать, коррелируют ли академические навыки участников с решением задач, поставленных в ходе эксперимента, В эксперименте участвуют студенты колледжа, поэтому исследователь решил уравнять две группы по среднему показателю успеваемости (СПУ). Для создания групп, эквивалентных по академическим способностям участников, которые отражены средним значением СПУ, необходимо выполнить следующие действия:

Шаг 1. Найдите значение переменной уравнивания для каждого испытуемого. В данном случае это сделать несложно, следует лишь получить данные о СПУ из деканата (с разрешения студентов, конечно). В других случаях для нахождения значения переменной уравнивания необходимо провести предварительное исследование участников, а это может потребовать дополнительного посещения испытуемыми лаборатории, что не всегда удобно (в этом заключается еще одна причина того, что исследователи любят использовать случайное распределение). Предположим, в эксперименте приняло участие 10 добровольцев (S), по 5 в каждой группе. Ниже приведены их СПУ:

S1:	3,24	S6:	2,45
S2:	3,91	S7:	3,85
S3:	2,71	S8:	3,12
S4:	2,05	S9:	2,91
S5:	2,62	S10:	2,21

Шаг 2. **Расположите СПУ** в порядке возрастания:

S4:	2,05	S9:	2,91
S10:	2,21	S8:	3,12
S6:	2,45	S1:	3,24
S5:	2,62	S7:	3,85
S3:	2,71	S2:	3,91

Шаг 3. **Разбейте все оценки** на **5** пар, в каждую из которых включите соседние **СПУ.**

Пара 1:	2,05 и 2,21
Пара 2:	2,45 и 2,62
Пара 3:	2,71 и 2,91
Пара 4:	3,12 и 3,24
Пара 5:	3,85 и 3,91

Шаг 4. Участников из каждой пары случайным образом распределите по группам: одного в группу 1, а второго в группу 2. Ниже показано одно из возможных распределений:

Группа 1	**Группа 2**
2,05	2,21
2,62	2,45
2,91	2,71
3,12	3,24

Окончание табл. 6.2

	3,85	3,91
Среднее арифметическое СПУ	2,91	2,90

Далее можно продолжить исследование с определенной уверенностью, что две группы эквивалентны друг другу по академическим способностям участников (2,91 — это практически то же, что и 2,90).

Примечание. Если исследуется более двух групп, процедура уравнивания остается без изменений до (включительно) шага 2. На шаге 3 вместо разбиения всех оценок по парам исследователь создает подгруппы, равные по количеству участников количеству экспериментальных групп. Далее, на шаге 4 участники каждой подгруппы случайным образом распределяются по группам.

Уравнивание нередко используется при небольшом количестве *(N)* участников, поскольку случайное распределение в таком случае ненадежно и может привести к созданию неэквивалентных групп. Однако для проведения уравнивания необходимо выполнение двух важных условий. Во-первых, вы должны быть уверены, что влияние переменной уравнивания на результаты эксперимента будут предсказуемым. Это означает, что переменная уравнивания должна коррелировать с зависимой переменной. Именно так обстояло дело в нашем гипотетическом исследовании памяти — тревожность несомненно снижала запоминание. При высокой корреляции переменной уравнивания и зависимой переменной статистические методы оценки планов позволяют выявить различия между группами. Если уравнивание проводится при низкой корреляции, вероятность обнаружить эти различия снижается. Это говорит о том, что необходимо быть очень внимательным при выборе переменной уравнивания.

Второе важное условие уравнивания заключается в том, что должен существовать адекватный способ измерения или определения значения переменной уравнивания для каждого участника. В некоторых исследованиях испытуемых предварительно тестируют для выявления значений переменной уравнивания, распределяют по группам, а затем проводят эксперимент. В зависимости от обстоятельств участникам может потребоваться два раза прийти в лабораторию, из-за чего могут возникнуть организационные проблемы. Предварительное тестирование может также создать у участников предварительное представление о целях исследования, что привдет к искажению. Проще всего провести уравнивание, если переменная уравнивания представляет собой конструкт, который можно определить без непосредственного исследования участников (например, значение среднего показателя успеваемости можно узнать из записей деканата), или если уравнивание проводится по независимой переменной. В исследовании памяти, к примеру, можно предварительно проверить память участников, далее уравнять их по результатам выполнения теста и распределить на две группы — с 2-секундной демонстрацией слов и 4-секундной. Таким образом их естественная способность к запоминанию будет находиться под контролем и различия в выполнении заданий будут связаны со скоростью демонстрации слов.

В главе 4 при обсуждении расслоенной выборки я отметил, что исследователи, использующие эту процедуру, сталкиваются с вопросом, сколько «слоев» необхо-

димо учитывать. Уравнивание ставит исследователя перед похожим вопросом. Нужно ли уравнивать группы по уровню тревожности в описанном выше исследовании памяти? А по уровню интеллекта? Или по уровню образования? Очевидно, что необходимо решить, какой показатель использовать, так как уравнивание трудно осуществить при наличии более одной переменной уравнивания. Из-за того что уравнивание становится невозможным, иногда приходится исключать некоторых испытуемых. Наличие проблемы выбора и измерения значений переменной уравнивания является одной из причин, по которой психологи зачастую предпочитают затратить силы на подбор достаточного количества добровольцев и провести случайное распределение, даже если они подозревают, что определенная внешняя переменная коррелирует с зависимой. Например, при исследовании памяти ученые редко задумываются об уровне тревожности, интеллекта или образования — они лишь набирают достаточно большие группы и допускают, что с помощью случайного распределения потенциальные осложнители будут равномерно распределены по экспериментальным ситуациям.

Внутрисубъектные планы

Как отмечалось в начале этой главы, при внутрисубъектном плане каждый испытуемый исследуется при каждом значении независимой переменной. Поскольку для всех участников данного вида исследования измерения проводятся несколько раз, иногда такой план называют планом с повторяющимися измерениями (например, см. главу 7). Одно практическое преимущество такого плана очевидно — для него требуется меньшее количество участников. Если вы проводите исследование, в котором сравниваются два условия, и хотите обследовать 20 испытуемых, для межсубъектного плана потребуется набрать 40 человек, а для внутрисубъектного — только 20.

Иногда внутрисубъектный план является единственным подходящим вариантом. В экспериментах по физиологической психологии, посвященных изучению ощущения или восприятия, часто сравниваются следствия условий, для создания которых не требуется много времени, но необходима большая предварительная подготовка. Например, в исследовании восприятия с помощью иллюзии Мюллера— Лайера проверить предположение, что иллюзия наиболее сильна при вертикальном расположении линий, можно, изменяя их направление (рис. 6.1). В качестве задания участников могут попросить с помощью определенной клавиши на клавиатуре компьютера изменять длину одной из отображаемых на экране линий и делать это до тех пор, пока линии не будут казаться одинаковыми. Каждая попытка может занять не более 5 секунд, а следовательно, абсурдно будет использовать переменную «расположение линий» в качестве межсубъектного фактора и исследовать разных людей по несколько секунд. Гораздо разумнее будет сделать переменную направления линий внутрисубъектым фактором и попросить участников выполнить последовательность заданий, охватывающих все значения переменной (а возможно, использующих каждое значение несколько раз). В отличие от исследования с «привлекательной/непривлекательной Барбарой» выполнение задания с одним условием не исключает участия в экспериментах с другими условиями.

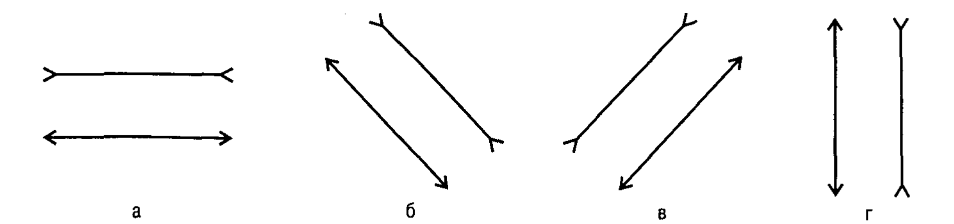


Рис. 6.1. Четыре примера иллюзии Мюллера-Лайера: а) горизонтальная, б) 45°, в) 135°, г) вертикальная

Кроме того, использование внутрисубъектного плана может потребоваться при небольшой численности исследуемой популяции и соответственно малом количестве добровольцев. В качестве примера можно привести исследование астронавтов или людей, обладающих каким-то особым мастерством (например, игроков в шахматы мирового класса). Несомненно, бывают случаи, когда даже при ограниченной популяции может возникнуть необходимость межсубъектного управления экспериментом. При оценке эффективности нового вида лечения редкой формы психопатологии может потребоваться сравнить проходящих лечение людей с людьми из контрольной группы, не подвергающейся экспериментальному воздействию.

Кроме практических удобств у использования внутрисубъектного плана есть еще одно преимущество — отсутствие проблемы эквивалентных групп, свойственной межсубъектным планам. Вспомните содержание главы 4, где рассказывалось о том, что в ходе заключительного статистического анализа результатов эксперимента для двух групп сравнивается изменчивость оценок между экспериментальными условиями с изменчивостью оценок в пределах каждого из условий. Изменчивость оценок между различными условиями может быть вызвана: а) независимой переменной, б) систематической дисперсией, вызванной осложнением, и/или в) несистематической дисперсией. В большой части дисперсия в межсубъектных планах вызывается индивидуальными различиями между испытуемыми в разных группах. При использовании внутрисубъектных планов дисперсия, связанная с индивидуальными различиями людей, исследуемых при разных условиях, не возникает. Рассмотрим конкретный пример.

Предположим, вы сравниваете дальность возможных бросков двух мячей для гольфа. Набрав 10 профессиональных гольферов, вы случайным образом распределяете их на две группы по 5 человек в каждой. Немного размявшись, каждый гольфер совершает удар по одному из двух мячей, что дает следующие результаты:

Профессионалы из первой группы	Мяч для гольфа № 1	Профессионалы из второй группы	Мяч для гольфа № 2
1	255	6	269
2	261	7	266
3	248	8	260
4	256	9	273
5	245	10	257
СА	253,00	СА	265,00
СО	6,44	СО	6,52

Следует отметить несколько особенностей. Во-первых, как показывает стандартное отклонение, в каждой группе наблюдается некоторая изменчивость оценок. Дисперсия в данном случае вызвана индивидуальными различиями в каждой группе, а также другими факторами. Во-вторых, заметно общее различие между группами: профессионалы из второй группы послали свой мяч дальше, чем профессионалы из первой группы. Почему это произошло? Возможны три причины:

1. Случайность: возможно, такое различие не является статистически значимым, а даже если и является, то остается вероятность 5%, что возникла ошибка 1-го рода (нулевая гипотеза истинна).
2. Мяч для гольфа: возможно, мяч для гольфа, используемый второй группой, способен укатиться на большее расстояние (в этом, конечно, и состоит гипотеза исследования).
3. Индивидуальные различия: гольферы во второй группе могли оказаться более сильными или способными.

Вероятность того, что все дело в индивидуальных различиях, снижена процедурой построения эквивалентных групп, описанной ранее. Используя случайное распределение или уравнивание, вы можете быть уверены, что вторая группа гольферов практически эквивалентна первой по способностям, силе и др. Но несмотря на это, возможно, что *некоторые* различия между группами могут быть вызваны индивидуальными различиями между членами этих групп. Во внутрисубъектных планах такая проблема просто не может возникнуть. Предположим, что вы повторяете исследование с участием только первых пяти гольферов и каждый из них сначала бьет по мячу № 1, а затем № *2.* В этом случае результаты будут следующими:

Профессионалы из первой группы	Мяч для гольфа № 1	Мяч для гольфа № 2
1	255	269
2	261	266
3	248	260
4	256	273
5	245	257
СА	253,00	265,00
СО	6,44	6,52

Тогда как для первого набора данных возможно три объяснения, для второго набора действительны только первые два объяснения. В первой таблице разница значений первого ряда (255 и 269) может быть вызвана случайностью, различиями мячей или индивидуальными различиями между профессионалами 1 и 6. Во втором наборе данных отсутствует вторая группа гольферов, а поэтому исключается возможность третьего объяснения. Таким образом, во внутрисубъектном плане из возможных предположений о причинах различий между экспериментальными условиями исключается предположение индивидуальных различий. Статистически это означает, что в случае внутрисубъектного плана заключительный анализ будет более чувствителен к небольшим различиям значений среднего арифметического, чем при межсубъектном плане.

Но постойте, достаточно ли вам, что во втором случае различия между первым и вторым наборами оценок могут быть вызваны *только:* а) случайными факторами и/или б) более высоким качеством второго мяча? Может быть, вы думаете, что первый гольфер каким-либо образом изменился между ударами по мячам № 1 и № 2? Хотя и маловероятно, что он успел нарастить 20 фунтов мускулов между двумя ударами, но, может быть, тут сыграл роль эффект тренировки или разминки? Или, возможно, гольфер заметил, что первый удар был не вполне правильным и исправил свою ошибку при ударе по второму мячу? А может быть, изменился ветер. Таким образом, основная проблема внутрисубъектного плана заключается в том, что после того, как испытуемый выполнил первую часть задания, приобретенный опыт или изменение обстоятельств может воздействовать на выполнение последующих его частей. Эта проблема может проявляться по-разному и называется **эффектом последовательности,** или **эффектом порядка.**

Прежде всего первая попытка может так подействовать на испытуемого, что вторая попытка будет выполнена лучше, как это происходит в случае эффекта тренировки. Кроме того, иногда повторение попыток постепенно приводит к усталости или скуке и задания выполняются все хуже и хуже. Эти два случая описывают действие **эффекта прогрессии.** Такое название используется, так как считается, что выполнение заданий равномерно (прогрессивно) изменяется от попытки к попытке. Определенные последовательности заданий могут также приводить к результату, отличному от того, который вызывается другими последовательностями. Такое явление называется **эффектом передачи.** Так, в исследовании с двумя базовыми условиями выполнение задания на условии *А* перед заданием на условие *В* может совсем иначе повлиять на испытуемого, чем выполнение задания на условии *В* перед заданием на условии *А.* Предположим, к примеру, что вы изучаете влияние шума на выполнение задания по сортировке карточек и используете для этого внутрисубъектный план. Участники должны рассортировать карточки по различным категориям за определенное количество времени. При условии *А* они должны проводить сортировку, когда их отвлекает шум в соседней комнате, возникающий в случайные моменты времени и потому непредсказуемый. При условии *В* слышно то же общее количество шума, но он распределен неслучайным образом и его появление можно предсказать. Если вы поместите людей сначала в условие *А,* а затем в условие *В,* то, скорее всего, они довольно плохо справятся с заданием на условии А (большинство людей показывают именно такой результат). Это может лишить их уверенности и повлиять на выполнение задания на условии *В.* Скорее всего, испытуемые будут лучше справляться с заданием *В,* но как только возникнет шум, они могут сказать себе: «Ну вот опять» и, возможно, перестанут стараться в достаточной мере. С другой стороны, если вы предложите в первую очередь задание *В,* в котором шум предсказуем, ваши добровольцы могут довольно хорошо справиться с ним (большинство людей справляется неплохо) и некоторая доля уверенности может перенестись и на вторую часть задания. Поэтому они могут лучше справиться с заданием *А,* чем вы ожидаете. Таким образом, задание на условии *А* может быть выполнено гораздо хуже в последовательности *А —В,* чем в последовательности *В— А.* Для условия *В* возникнет та же проблема. Говоря ко-

ротко, последовательность предъявления заданий вне зависимости от эффекта тренировки или усталости может повлиять на результаты исследования. В исследованиях, где можно ожидать появления эффекта передачи, экспериментаторы часто используют межсубъекный план. И действительно, в исследованиях по сравнению предсказуемого и непредсказуемого шума участников обычно разделяют на две группы. Однако использование межсубъектного плана не всегда приемлемо, и поэтому очень важно уметь контролировать эффект последовательности во внутрисубъектном плане.

Проблема контроля за эффектом последовательности

Обычно эффект последовательности контролируется с помощью создания нескольких последовательностей — такой подход известен как **позиционное уравнивание.** Как вы узнаете далее, данная процедура лучше справляется с устранением эффекта прогрессии, чем эффекта передачи. В зависимости от того, исследуются ли участники один раз или более при каждом из экспериментальных условий, различают два основных вида позиционного уравнивания.

Однократное исследование при каждом наборе условий

В некоторых экспериментах испытуемые исследуются только один раз при каждом условии. В качестве примера рассмотрим интересное исследование Рейнолдса, посвященное способности игроков в шахматы определять уровень мастерства других игроков, которое было проведено в 1992 г. В различных шахматных клубах Нью-Йорка Рейнолдс набрал 15 игроков с разным уровнем мастерства и попросил их рассмотреть шесть партий, в которых игра еще только развивалась (т. е. от ее начала было сделано около 20 ходов). При каждой попытке игроки изучали доску с незавершенной партией (им сказали, что пары игроков в каждой игре имели равные способности) и оценивали уровень развития навыков игроков в соответствии со стандартной системой оценок. Расположение фигур в каждой партии отражало различные уровни мастерства игроков. Рейнолдс обнаружил, что игроки, обладающие наибольшими способностями, точнее других определяли уровень мастерства, отраженный в изучаемой ими партии.

Очевидно, что в исследовании Рейнолдса использовалась внутрисубъектная переменная — каждый из 15 участников изучал все шесть партий. Также вполне естественно, что игроки изучали каждую партию только один раз. Вследствие этого, Рейнолдс столкнулся с проблемой контроля за возможным эффектом последовательности. Он безусловно не хотел, чтобы все 15 участников рассматривали партии в одинаковой последовательности, но как можно было избежать этого?

Завершенное позиционное уравнивание

Если во внутрисубъектном плане участники исследуются один раз при каждом условии, решить проблему последовательности может **полное позиционное уравнивание.** Это означает, что все возможные последовательности будут использова-

ны хотя бы один раз. Чтобы найти общее количество требуемых последовательностей, необходимо вычислить $X!$, где X — это количество условий, а «!» — знак факториала. Например, если в исследовании используются три условия, то можно создать шесть последовательностей:

$$3! = 3 \times 2 \times 1 = 6.$$

Эти последовательности для исследования с условиями *A, B и C* будут следующими:

ABC *BAC*

ACB *CAB*

BCA *CBA*

Проблема полного позиционного уравнивания состоит в том, что по мере увеличения количества условий количество необходимых последовательностей растет экспоненциально. Для трех условий требуется 6 последовательностей, но увеличение количества условий всего до четырех дает необходимость построения 24 последовательностей ($4 \times 3 \times 2 \times 1$). Как вы можете догадаться, в случае исследования Рейнолдса полное позиционное уравнивание было возможно только при наборе гораздо большего количества участников, чем набранные им 15 человек. По сути, при использовании шести партий (т. е. условий), чтобы охватить все возможные последовательности, ему потребовалось бы найти 6!, или 720 игроков в шахматы. Очевидно, что в этом случае использовался другой подход.

Частичное позиционное уравнивание

Использование подмножества от общего количества последовательностей дает нам **частичное позиционное уравнивание.** Именно в этом и состояло решение Рейнолдса — он позаботился о том, чтобы «последовательность демонстрации была случайной для каждого субъекта» (Reynolds, 1992, p. 411), а тем самым просто сделал случайную выборку из 720 возможных последовательностей. Выборки из набора последовательностей часто используются в ситуациях, когда количество участников меньше количества возможных последовательностей или при большом числе условий.

Рейнолдс сделал выборку из общего набора последовательностей, но также можно было использовать и другой широко применяемый метод — правильный **латинский квадрат.** Этот метод получил свое имя от древней римской загадки о том, как расположить в матрице латинские буквы так, чтобы каждая буква встречалась в каждом ряду и каждом столбце только один раз (Kirk, 1968). Построить латинский квадрат сложнее, чем выбрать случайное подмножество из целого, но, построив его, вы можете быть уверены, что а) частота появления каждого экспериментального условия одинакова для всех последовательных позиций и б) каждому условию предшествует, а также следует за ним каждое другое условие строго один раз. В табл. 6.3 показано, как построить латинский квадрат размером 6×6. Каждую из букв примите за одну из шести партий, изучаемых игроками в исследовании Рейнолдса.

Таблица 6.3

Построение правильного латинского квадрата

В правильном латинском квадрате каждое экспериментальное условие в каждой последовательной позиции встречается одинаково часто и каждому условию предшествует, а также следует за ним каждое другое условие строго один раз. Ниже показано, как построить квадрат размером 6x6.

Шаг 1. Постройте первый ряд в соответствии со следующим правилом:

A B «X» C «X–1» D «X–2» E «X–3» F и т. д.,

где А означает первое экспериментальное условие, а «*X*» — последнее. Для построения квадрата размером 6 × 6 в первом ряду будут сделаны следующие замены:

X = шестая буква алфавита → *F*;

X–1 = пятая буква → E.

Таким образом, первый ряд будет состоять из следующих букв:

A B ***F*** (заменяя «X») *C* ***E*** (заменяя «*X*–1») *D*.

Шаг 2. Постройте второй ряд. Прямо под каждой буквой первого ряда во втором ряду поместите следующую по алфавиту букву, единственное исключение — буква *F*. Дойдя до нее, вернитесь к началу алфавита и поместите под ней букву *A*. Получится следующее:

A B F C E D
B C A D F E

Шаг 3. Постройте оставшиеся четыре ряда следуя правилу, изложенному на шаге 2. Таким образом, конечный квадрат размером 6x6 будет:

A B F C E D
B C A D F E
C D B E A F
D E C F B A
E F D A C B
F A E B D C

Шаг 4. Чтобы задать действительную последовательность условий для каждого ряда, случайным образом поставьте в соответствие буквам *от A* до Fшесть экспериментальных условий. Для каждого ряда выделите одинаковое количество участников.

A	*B*	*F*	*C*	*E*	*D*
B	*C*	***A***	*D*	*F*	*E*
C	*D*	*B*	*E*	***A***	*F*
D	*E*	*C*	*F*	*B*	***A***
EF	*D*	***A***	*C*	*B*	
*F**A***	*E*	*B*	*D*	*C*	

Я выделил условие *A* (партию *A),* чтобы показать вам, как квадрат выполняет два условия, указанные в предыдущем абзаце. Во-первых, условие *A* появляется в каждой из шести последовательных позиций (первым в первом ряду, третьим во втором и т. д.). Во-вторых, за Л каждая другая буква следует строго один раз. В направлении от верхнего ряда к нижнему (1) за *A* каждая другая буква следует строго один

раз. В направлении от верхнего ряда к нижнему (1) за А следуют: *B, D, F,* ничего и *Еп* (2) А предшествуют: ничто, *С, Е,В,Du F.* Это верно **и** для всех остальных букв. Чтобы использовать латинский квадрат размером 6х6, необходимо случайным образом поставить в соответствие каждому из шести экспериментальных условий (шесть различных шахматных партий у Рейнолдса) одну из шести букв, от *А* до *F.*

При использовании латинского квадрата необходимо, чтобы количество участников равнялось или было кратно количеству рядов квадрата. То, что в исследовании Рейнолдса было 15 участников, указывает, что он не использовал латинский квадрат. Если бы он добавил еще трех игроков, получив в целом 18, он мог бы случайным образом распределить их по шести радам квадрата (3x6 = 18).

Многократное исследование при каждом наборе условий

В исследовании Рейнолдса не было особой причины просить игроков рассматривать каждую партию более одного раза. И в случае экспериментов с памятью, когда участники запоминают и воспроизводят наборы слов в порядке, заданном латинским квадратом размером 4 х 4, им редко приходится проделывать это дважды. Экспериментатор может попросить участников повторять задания, только если он изучает, как повторение влияет на запоминание. Однако для многих исследований выполнение заданий испытуемыми более одного раза при каждом наборе условий не только целесообразно, но и необходимо. Это может потребоваться, к примеру, при исследовании ощущения или восприятия. Пример такого исследования показан на рис. 6.1.

Предположим, вы проводите исследование и хотите проверить, действительно ли иллюзия возникает в большей степени, если рисунок предъявляют вертикально, а не горизонтально или под углом в 45°. Четырем условиям исследования соответствуют четыре буквы алфавита:

А = горизонтально.
В = 45° влево.
С = 45° вправо.
D = вертикально.

Участники исследования видят рисунок на экране компьютера. Они должны изменять длину параллельных линий пока те не будут выглядеть одинаковыми. Есть две базовые процедуры, с помощью которых можно представить участникам четыре экспериментальных условия.

Обратное позиционное уравнивание

При использовании **обратного позиционного уравнивания** экспериментатор представляет условия в определенном порядке, а затем делает это еще раз, изменяя порядок на противоположный. В эксперименте с иллюзией порядок может быть следующий: A-B-C-D, а затем D-C-B-A. Если исследователь хочет, чтобы для каждого из условий участники выполняли задание более двух раз, что нередко бывает при исследовании восприятия, последовательность должна повторяться столько раз, сколько необходимо. Так, если вы хотите, чтобы испытуемые рассматривали каждую из четырех иллюзий, приведенных на рис. 6.1, по восемь раз и используете

для этого обратное позиционное уравнивание, испытуемые получат следующую последовательность заданий:

A-B-C-D - D-C-B-A - A-B-C-D - D-C-B-A - A-B-C-D - D-C-B-A -
-A-B-C-D-D-C-B-A.

Обратное позиционное уравнивание было использовано в одном из самых известных психологических исследований, проведенном в 30-х гг. XX в. Дж. Ридли Струпом. Возможно, вы и сами проходили тестирование по методике Струпа — в нем показывают названия цветов, напечатанные разными чернилами, причем названия не соответствуют цвету чернил, и просят называть цвета, не читая названий. В случае если показывают напечатанное синими чернилами слово «КРАСНЫЙ», правильный ответ — «синий», а не «красный». Исследование Струпа — это классический пример одного из видов экспериментального плана, описанного в следующей главе, поэтому более подробно о работе ученого вы узнаете, изучив вставку 7.1[1].

Блоковая рандомизация

Второй способ представления последовательности условий, когда каждое из них исследуется более одного раза, — это **блоковая рандомизация.** Об этой процедуре рассказывалось ранее при описании случайного распределения участников по группам в межсубъектных экспериментах. Основное правило блоковой рандомизации состоит в том, что все условия используются по одному разу, прежде чем любое из них встречается во второй раз. В пределах каждого блока порядок расположения условий случайный, что предотвращает возможность предугадывания участниками хода событий (проблема, которая может возникнуть при обратном позиционном уравнивании).

В примере с иллюзией (рис. 6.1) участники в случайном порядке встречаются со всеми четырьмя условиями, затем еще раз, но уже в блоке с другим случайным порядком их следования, и т. д. столько раз, сколько требуется. Порядок условий, полученный обратным позиционным уравниванием будет следующим:

A-B-C-D D-C-B-A.

С помощью блоковой рандомизации, кроме прочих, можно получить одну из двух последовательностей:

B-C-D-A C-A-B-D или C-A-B-D A-B-C-D.

Чтобы получить представление о том, как в ходе реального внутрисубъектного эксперимента проводится блоковая рандомизация, рассмотрим исследование слухового восприятия, проведенное Карелло с соавторами (Carello, Anderson & Kunkler-Peck, 1988).

[1] Хотя обратное позиционное уравнивание обычно используется, когда участники исследуются более одного раза при каждом условии, его принципы можно применить и к внутрисубъектным планам, в которых каждый испытуемый встречается с каждым из условий только один раз. Так, например, если во внутрисубъектном исследовании используется шесть различных условий, и каждое изучается один раз для каждого испытуемого, то половина участников получит последовательность A-B-C-D-E-F, а другая половина - обратную последовательность (F-E-D-C-B-A).

Пример 5. Позиционное уравнивание с помощью блоковой рандомизации

Способность людей определять местонахождение источника звуков известна с давних пор: в обычных условиях мы довольно легко определяем направление, откуда приходит звук. Карелло и ее исследовательская группа заинтересовались вопросом, могут ли люди определить размер объекта по звуку, издаваемому им при падении на пол. Для изучения данного вопроса был разработан аппарат, изображенный на рис. 6.2. Испытуемые слышали, как деревянная рейка падала на пол, и пытались определить ее длину. В качестве ответа участники за одну попытку изменяли расстояние между краем стола, за которым они сидели, и подвижной вертикальной поверхностью. Попыткой называлось падение одной и той же рейки с определенной высоты пять раз подряд. Слушая звук падений, участники передвигали стенку вперед и назад, пока расстояние между ней и столом не становилось равным длине рейки. В первом из двух экспериментов внутрисубъектной независимой переменной была длина рейки. Она принимала семь значений: 30,45,60,75, 90,105 и 120 см. Каждый участник определял длину каждой рейки три раза. Как и следовало ожидать, экспериментаторы провели позиционное уравнивание последовательности длин реек и сделали это с помощью блоковой рандомизации. Таким образом, семь реек разной длины использовались в одном случайном порядке, затем в другом и наконец в третьем. Еще раз обратите внимание на важную особенность блоковой рандомизации: каждая длина использовалась по одному разу, прежде чем одна из них использовалась повторно, и по два раза, прежде чем какая-либо из них была использована в третий раз.

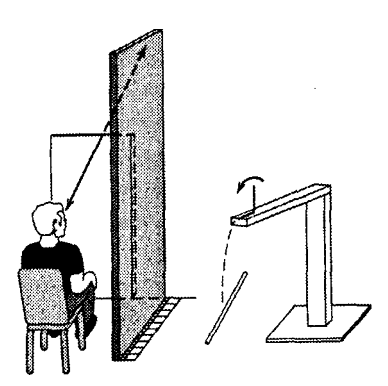


Рис. 6.2. Схема экспериментального устройства (Carello, Anderson, and Kunkler-Peck, 1998). Услышав звук падения рейки, испытуемый изменял расстояние между краем стола, за которым сидел, и обращенной к нему лицом вертикальной поверхностью так, чтобы оно равнялось длине рейки

Хотя задание на определение длины рейки может показаться очень сложным, испытуемые справились с ним на удивление легко. Результаты второго эксперимента, который повторял первый, но с более короткими рейками (от 10 до 40 см), были даже выше, чем результаты первого (рейки от 30 до 120 см). Стоит отметить еще две особенности этих экспериментов. Во-первых, они представляют собой хороший пример стратегии, обычно используемой при исследовании восприятия: внутрисубъектный план, который требует небольшого количества испытуемых

и большого числа попыток. В первом эксперименте участвовали восемь студентов, а во втором — шесть, причем каждый из них выполнял задание 21 раз. Во-вторых, вспомните пробные исследования, которые обсуждались в главе 3. Цель таких исследований — испытание экспериментальной процедуры и изменение ее в случае обнаружения затруднений. Нечто подобное произошло и в обсуждаемом нами исследовании, хотя первый эксперимент в действительности не был пробным. Одна из особенностей процедуры первого эксперимента привела к изменениям во втором эксперименте. В первом случае рейки падали на пол, а во втором — на приподнятую поверхность. Но почему? По словам Карелло, Андерсон и Канкер-Пек, причина была весьма практической (читая между строк, заметим, что аспиранты, собирающие данные, были очень благодарны устранению этого фактора): изменение процедуры помогло «снизить нагрузку на колени и спину экспериментаторов» (р. 212).

Проблемы процедуры позиционного уравнивания

Позиционное уравнивание не всегда помогает снизить эффект последовательности, так как оно рассчитано только на линейный эффект. Во многих случаях, а особенно в случае эффекта передачи, он нелинеен. Чтобы понять, почему это вызывает проблемы, рассмотрим следующий гипотетический пример.

Предположим, вы проводите эксперимент с участием людей и сравниваете два вида лабиринтов, аналогичных тем, которые показаны на рис. 6.3. Один из них *(А)* последовательный и предполагает серию поворотов направо или налево, а второй *(В)* имеет более сложную структуру и напоминает лабиринт Хэмптон Корт. В исследовании с такими лабиринтами испытуемым завязывают глаза и просят найти выход, для определения направления поворотов используя карандаш или указку.

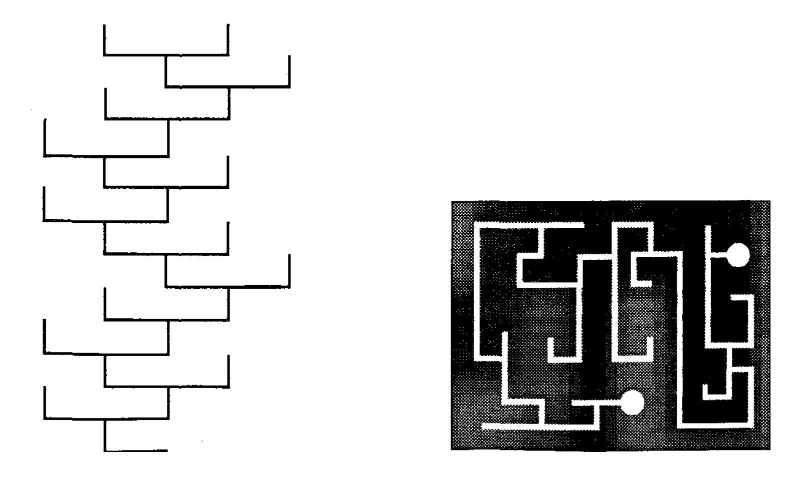


Рис. 6.3. Два типа лабиринтов, которые люди проходят без зрительного контроля: а) последовательный лабиринт и б) пространственный лабиринт

Допустим, что в исследовании используется внутрисубъектный план. Половина участников изучают лабиринт Л, а затем *В,* а другая половина — *В,* а затем Л, что дает полное позиционное уравнивание условий. Предположим, сеанс длится один час и со временем участники устают или начинают испытывать скуку, а следовательно, результаты прохождения второго лабиринта могут снизиться. Разумно предположить, что нарастание усталости в течение этого часа будет проходить в соответствии с эффектом прогрессии, т. е. она будет изменяться *линейно* от попытки к попытке. Поэтому позиционное уравнивание, обеспечивающее, чтобы каждый лабиринт исследовался одинаковое количество раз первым и вторым по счету, уравновесит действие усталости. Допустим, что усталость добавляет три ошибки к общим оценкам и что лабиринт *В* (в котором в среднем совершается 15 ошибок) сложнее, чем лабиринт Л (со средним количеством ошибок, равным 10). Для последовательностей *А* > *В* и *В* > *А* возможны следующие оценки:

	Ошибки вызваны		Общее количество ошибок
	сложностью лабиринта	скукой	
Лабиринт А, а затем	10	0	10
лабиринт *В*	15	+3	18
Лабиринт *В,* а затем	15	0	15
лабиринт А	10	+3	13

Объединение этих последовательностей приведет к тому, что усталость будет одинаково влиять на прохождение обоих лабиринтов, и поэтому действие ее сотрется. Среднее количество ошибок составляет 11,5 для лабиринта Л [(10 + 13)/2] и 16,5 для более сложного лабиринта *В* [(18 + 15)/2].

Но, как отмечалось ранее, эффект передачи может вызвать проблемы, с которыми позиционное уравнивание не сможет справиться. Предположим, к примеру, что решение лабиринта Л поможет людям понять, как в принципе решать лабиринты, а решение лабиринта Б не *приведет* к такому пониманию. В таком случае в последовательности Л > *В* изучение первым лабиринта *А* повлечет за собой перенос знаний на лабиринт *В,* тогда как в последовательности *В* > *А* изучение первым лабиринта *В* не приведет к положительному переносу. Другими словами, две последовательности будут иметь **асимметричный перенос** (Poulton, 1982). Это означает, что одна из них дает результаты, которые невозможно уравнять с помощью позиционного уравнивания. Предположим, что в примере с лабиринтами изучение первым лабиринта Л приводит к тому, что лабиринт *В* становится очень просто изучить, а именно, это снижает общее количество ошибок для лабиринта *В* на 10. При этом изучение первым лабиринта *В* не приводит к переносу результатов на лабиринт Л. Таким образом:

	Ошибки вызваны			Общее количество ошибок
	сложностью лабиринта	переносом	скукой	
Лабиринт *А,* а затем	10			10
лабиринт *В*	15	-10	+3	8
Лабиринт В, а затем	15			15
лабиринт А	10	0	+3	13

К

Объединение двух последовательностей приводит к тому, что действие усталости уравнивается, а влияние асимметричного переноса — нет. Среднее количество ошибок составляет 11,5 для обоих лабиринтов: для лабиринта *A* [(10+ 13)/2] и для предположительно более сложного лабиринта *B* [(8 + 15)/2]. Проблема переноса приведет к тому, что между двумя лабиринтами не будет обнаружено никаких различий, что для исследователя явится весьма неприятным сюрпризом. При подозрении на асимметричный перенос стоит, если возможно, перейти к межсубъектному плану.

Проблемы контроля в исследованиях по психологии развития

Как вы уже знаете, чтобы решить, какой план использовать, внутрисубъектный или межсубъектный, исследователю необходимо учитывать различные факторы. Иногда в исследованиях по психологии развития требуется дополнительное рассмотрение данного вопроса. Это касается использования метода поперечных срезов и лонгитюдных исследований.

Если вы изучали психологию развития или детскую психологию, то могли уже встречаться с этими терминами. Основной переменной таких исследований является возраст, ведь основная задача психологии развития состоит в том, чтобы определить, как мы меняемся с возрастом. **Исследование методом поперечных срезов** является межсубъектным. При сравнении речевой деятельности 3-, 4- и 5-летних детей методом поперечных срезов будут изучаться различные группы детей. В лонгитюдном исследовании одна группа изучается на протяжении некоторого отрезка времени. Оно является внутрисубъектным, или исследованием с повторяющимися измерениями. При аналогичном исследовании речи будет изучаться речевая деятельность 3-летних детей, а затем те же дети будут обследоваться, когда им исполнится 4 года и 5 лет.

Явное преимущество метода поперечных срезов в случае эксперимента с речью — это экономия времени. Такое исследование может занять около месяца, а лонгитюдное исследование того же вопроса займет 3 года. Однако в некоторых исследованиях, проводимых методом поперечных срезов, возникает особый вид проблемы неэквивалентных групп, известный как **эффект когорты.** Когорта — это группа людей, рожденных приблизительно в одно время. Если вы исследуете три возрастные группы, то они различаются не только по возрасту, но и по условиям, в которых они росли. Эта проблема не очень заметна при изучении 3-, 4- и 5-летних детей, но что будет, если вы захотите исследовать, снижается ли интеллект с возрастом, и решите сравнить группы людей возрастом 30,50 и 70 лет? Вы действительно можете обнаружить снижение интеллекта с возрастом, но будет ли это действительно означать постепенное снижение или разница будет связана с различными условиями жизни этих трех групп? Например, 70-летние пошли в школу во время Великой депрессии, 50-летние учились во время подъема после Второй мировой войны, а 30-летние росли глядя в телевизор. Эти факты могли исказить результаты. И действительно, подобные искажения имели место. В ранних исследованиях по изучению влияния возраста на 7(5было обнаружено заметное снижение

интеллекта, но эти исследования проводились методом поперечных срезов (например, Miles, 1933). Более поздние лонгитюдные исследования показали другую зависимость (Schaie, 1988). Например, вербальные способности снижаются очень мало, особенно если человек сохраняет речевую активность (мораль: используй или потеряешь).

Эффект когорты может повредить исследованиям, проводимым методом поперечных срезов, но у лонгитюдных исследований также есть свои проблемы, в частности связанные с истощением. Если большое количество испытуемых перестанет участвовать в исследовании, завершающая эксперимент группа будет сильно отличаться от набранной вначале. В примере с исследованием взаимосвязи возраста и *IQ* здоровые люди могут вести более интеллектуальную жизнь, чем постоянно болеющие. Испытуемые, страдающие серьезными хроническими заболеваниями, могут умереть до завершения исследования, а следовательно, оставшаяся группа будет иметь более высокий уровень интеллекта, чем та, с которой проводились первые эксперименты. Могут также возникнуть этические проблемы. По мере развития и взросления люди могут изменить свое отношение к участию в исследовании. Большинство ученых, занимающихся лонгитюдными исследованиями, пришли к выводу, что получение обоснованного согласия — это не единичное событие, а процесс. Исследователи, обращающие пристальное внимание на этический аспект своей работы, при проведении длительных исследований периодически (раз в несколько лет) повторно получают согласие испытуемых (Fischman, 2000).

Стараясь снизить эффект когорты и решить проблему истощения, некоторые исследователи комбинируют метод поперечных срезов и лонгитюдный подход. Например, К. Уорнер Шаи в 1983 г. опубликовал результаты ряда экспериментов, известных как «Сиэтлское лонгитюдное исследование», в ходе которых каждые семь лет (начиная с 1950 г.) набиралась и обследовалась новая когорта. Использование групп с разницей в возрасте в семь лет ослабило эффект когорты, а добавление новых участников через определенные промежутки времени уменьшило воздействие истощения. Любая когорта может испытать сильное истощение, но если когорт много, то хотя бы несколько из них долгое время останутся относительно невредимыми.

Длительность сиэтлского проекта впечатляет, но мировой рекорд среди исследований с повторяющимися измерениями принадлежит, вероятно, самому известному лонгитюдному исследованию — исследованию одаренных детей Льюиса Термана. Вставка 6.1 представляет хронику этой работы.

Проблемы искажения

Поскольку в психологических исследованиях экспериментаторы, а обычно и испытуемые, — это люди, то есть вероятность появления некоторого «искажения», попытки предугадать, что должно произойти в ходе эксперимента, и ожидания этих событий, способного повлиять на результаты этих исследований. Есть разные виды искажений, но все их можно разделить на две большие группы: связанные с экспериментаторами или вызываемые испытуемыми. Эти два вида искажений нередко переплетаются друг с другом.

ВСТАВКА 6.1

Классические исследования — самое долгое из повторяющихся измерений

В 1921 г. психолог Льюис Терман (1877-1956) начал исследование, ставшее самым долгим исследованием с повторяющимися измерениями в истории психологии. Терман сам был не по годам развитым ребенком и всегда интересовался изучением одаренных детей. Его докторская диссертация, написанная под руководством Эдмунда Сэнфорда в Университете Кларк, которую он защитил в 1905 г., была его первым серьезным исследованием одаренности. В своей работе он изучал сообразительных и отстающих учеников местной школы и хотел определить, какие тесты покажут различия между ними наилучшим образом (Minton, 1987). Ранняя заинтересованность одаренностью и проверкой умственных способностей явилась предвестницей того вклада, который Терман сделал в психологию. На основе теста интеллекта, созданного французом Альфредом Бине, он разработал до сих пор популярный тест интеллекта Сэнфорд-Бине, а также начал лонгитюдное исследование одаренных детей, которое продолжалось еще долго после его смерти.

Терман, как и большинство ученых того времени, занимавшихся тестированием интеллекта, был убежден, что Америка должна стать страной, в которой положение человека будет определяться его способностями, а у власти будут находиться те, кто имеет наибольшие *способности* к управлению. Вполне понятно, как такое убеждение привело его к изучению *10* и одаренности. Чтобы создать такое общество, необходимо найти способы определения (т. е. измерения) таланта и разработать подходы к его развитию.

В своей диссертации Терман изучал всего 14 детей, но его лонгитюдное исследование одаренных детей приняло гигантские размеры. С помощью различных процедур отбора он набрал 1470 детей (824 мальчика и 646 девочек), большинство из которых учились в начальной школе, а 444 - в средней или старшей (цифры взяты из Minton, 1988). Средний коэффициент интеллекта у детей был 150, а это значит, что группа принадлежала к самой верхушке интеллектуальной элиты популяции. Терман набрал группу аспирантов, и с их помощью каждому ребенку была выдана развернутая батарея тестов и опросников. К моменту завершения начального этапа тестирования дело каждого ребенка составило около 100 страниц (Minton, 1988)! Результаты первого анализа группы были опубликованы в работе «Умственные и физические особенности тысячи одаренных детей» (Terman, 1925).

Терман собирался провести лишь небольшое повторное исследование, но проект занял всю его жизнь. Выборка была заново протестирована в конце 20-х гг. (Burks, Jensen, & Terman, 1930), а результаты дополнительных исследований, проведенных Терманом за время его жизни, были опубликованы через 25 (Terman & Oden, 1947) и 35 (Terman & Oden, 1959) лет после первого тестирования. После смерти Термана руководство проектом принял Роберт Сире, известный психолог. Во введении к работе, опубликованной через 35 лет после первого тестирования, Сире писал: «С учетом всех данных существует немалая вероятность того, что последние результаты исследования последнего одаренного ребенка Термана будут получены не ранее 2010 года!» (Terman & Oden, 1959, с. ix). Между 1960 и 1986 г. Сире провел пять дополнительных исследований и работал над обширным исследованием участников группы пожилого возраста, когда смерть прервала его работу в 1989 г. (Cronbach, Hastorf, Hilgard, & Maccoby, 1990). Его книга, названная «Одаренная группа в преклонном возрасте», была опубликована в 1995 г. (Holahan, Sears & Cronbach, 1995).

Следует отметить три особенности этого металонгитюдного исследования. Во-первых, работа Термана разрушила стереотипное представление об одаренных детях как о выдающихся,

но социально отсталых и, как правило, «рано сгорающих». Члены его группы выделялись своими способностями и были хорошо приспособлены к жизни, а в дальнейшем, когда выросли, смогли добиться успеха. К зрелому возрасту «группа произвела на свет тысячи научных статей, 60 документальных книг, 33 романа, 375 рассказов, 230 патентов, а также многочисленные теле- и радиопередачи, произведения искусства и музыкальные произведения» (Hothersall, 1990, p. 353). Во-вторых, собранные коллективом Термана данные остаются богатым архивным источником информации для современных исследователей. Например, были опубликованы исследования профессиональных достижений одаренных семей из группы Термана (Tomlinson-Keasy, 1990) и свидетельств, предвещающих долголетие в этой группе (Friedman et al., 1995). В третьих, если вспомнить об истощении, неизбежно настигающем лонгитюдные исследования, повторные исследования Термана потрясают с методологической точки зрения. Следующие цифры (взяты из Minton, 1988) показывают процент испытуемых, принявших участие в первых трех исследованиях, от общего количества оставшихся в живых участников:

Через 10 лет: 92%.
Через 25 лет: 98%.
Через 35 лет: 93%.

Цифры очень высоки и говорят о преданности друг другу Термана и его группы. Члены группы называли себя «термитами» и некоторые даже носили украшения в форме термитов (Hothersall, 1990). Терман переписывался с сотнями своих испытуемых и искренне заботился о них. Ведь группа состояла из людей, которые, по мнению Термана, были ключом к будущему Америки.

Искажение, вызываемое экспериментатором

Влияние **искажения, вызываемого экспериментатором,** на результаты исследования часто показывают на примере истории Умного Ганса (вставка 3.3). Дрессировщик знал ответ на вопрос «Сколько будет 3 умножить на 3?» и, незаметно кивая головой, подавал сигналы, понятные для «умной» лошади. При проверке гипотезы также действия экспериментатора могут подтолкнуть участников вести себя так, чтобы подтвердить сделанные исследователем предположения. Хотя стереотип ученого подразумевает абсолютную объективность, бесстрастность и даже механистичность, мы знаем, что исследователи довольно эмоционально относятся к своей работе. Легко представить, как желание подтвердить гипотезу, которой строго придерживаются, может привести к тому, что поведение ученого повлияет на результаты исследования.

Например, предубежденный экспериментатор может при разных условиях по-разному воздействовать на испытуемых. Роберт Розенталь разработал процедуру, демонстрирующую этот факт. Участникам нескольких из его исследований (например, Rosenthal & Fode, 1963a) показывали фотографии лиц и просили высказать свое мнение по поводу изображенных людей. Например, испытуемых просили разложить фотографии в ряд по признаку преуспевания людей с использованием ранговой шкалы с делениями от -10 (неудачник) до +10 (весьма преуспевающий). Все участники получали одни и те же фотографии и выполняли одинаковые задания. Независимой переменной было ожидание экспериментатора. Одним экспериментаторам сказали, что большинство людей в случае сомнения решат дело в по-

ложительную сторону и позитивно оценят изображенных людей, а другим советовали ожидать негативных оценок. Любопытно, что ожидание экспериментатора обычно влияло на то, как участники раскладывали фотографии, даже при условии, что они были одинаковыми для обеих групп. В чем причина этого явления?

Согласно Розенталю (Rosenthal, 1966), экспериментаторы невольно и неявным образом передавали свои ожидания. К примеру, пока испытуемый оценивал изображение, экспериментатор держал фотографию в руках. Если экспериментатор ожидал «+8» а участник говорил «-9», то как первый мог отреагировать? Возможно, слегка нахмуриться? Как мог участник ответить на то, что экспериментатор нахмурился? Может быть, попробовать в следующий раз сказать «+7» и посмотреть, вызовет ли это улыбку или кивок экспериментатора? Возможно ли, что экспериментатор в такой ситуации, сам того не осознавая, незаметно влиял на ответы испытуемых? Разве это не напоминает вам историю с Умным Гансом?

Розенталь смог даже показать, что ожидания экспериментаторов могут передаваться объектам в исследованиях с животными. Например, крысы быстрее изучают лабиринт, если экспериментатор *думает,* что они были специально выведены для этого, чем если он считает их не способными к прохождению лабиринта (Rosenthal & Fode, 1963b). Крысы, конечно, случайным образом разделены между экспериментаторами и имеют одинаковые способности. Ситуация, возможно, объясняется тем, что экспериментаторы, думая, что их крысы обладают великолепной способностью проходить лабиринты, лучше с ними обращаются — например, больше гладят, что влияет на обучение.

Следует отметить, что статистический аспект исследования Розенталя был подвергнут критике. Его работу также осуждали за то, что результаты интерпретировались как вызванные ожиданием, тогда как на них могли повлиять и другие факторы. Например, Барбер (Barber, 1976) поднял вопрос о валидности некоторых статистических выводов, сделанных в работе Розенталя. Согласно Барберу, по крайней мере в одном из исследований 3 из 20 экспериментаторов получили результаты противоположные созданным у них ожиданиям. Розенталь не включил этих экспериментаторов в анализ и получил существенные различия между результатами остальных 17 человек. Однако при учете в ходе анализа всех 20 экспериментаторов эти различия исчезли. Также Барбер указал, что в анализах исследований с животными некоторые результаты объясняются тем, что экспериментаторы просто подделывали данные (например, не записывали ошибки в прохождении лабиринта). Другая проблема исследования Розенталя состоит в том, что его процедура не совпадает с той, что обычно используется в экспериментах, ведь большинство экспериментаторов исследуют всех участников при всех экспериментальных условиях, а не только выполнение задания при одном из условий. Таким образом, Розенталь мог переоценить степень возникшего искажения.

Несмотря на возникшие сомнения, нельзя исключить эффект ожидания со стороны экспериментатора — исследование Розенталя было воспроизведено не только им и его коллегами, но также и другими исследователями в различных условиях (например, Word, Zanna & Cooper, 1974). Кроме того, экспериментатор может влиять на результаты исследования не только своими определенными ожиданиями. Поведение участников может изменяться под влиянием расы или пола иссле-

дователя, а также его манеры поведения, дружелюбности и общего отношения к испытуемым (Adair, 1973). В качестве примера можно привести исследование, проведенное Фрейсс и Деспрелс-Фрейсс (Fraysse and Desprels-Fraysse, 1990), в ходе которого было обнаружено, что отношение экспериментатора может повлиять на выполнение дошкольниками заданий на классификацию. Под руководством «заботливого» экспериментатора дети заметно лучше справлялись с заданием, чем в случае с «безразличным» исследователем.

Контроль за искажением, вносимым экспериментатором

Полностью устранить влияние экспериментатора невозможно, так как нельзя превратить его в машину, но можно механизировать процедуру исследования. В этом и состоит один из способов уменьшения искажения. К примеру, совсем не сложно отстранить нахмуренного или улыбающегося исследователя от проведения эксперимента на восприятие. С помощью современных компьютерных технологий можно показывать фотографии на экране и попросить испытуемых отвечать нажатием клавиш. Экспериментатор при этом может находиться в соседней комнате.

Процедуры автоматизированного исследования животных применяются с 20-х гг. XX в. и позволяют полностью устранить воздействие человека. Е. К. Толмен, не дожидаясь появления компьютеров, изобрел «самозаписывающий лабиринт со столом для автоматической доставки» (Tolman, Tryon, & Jeffries, 1929). «Стол для автоматической доставки» получил свое название, поскольку он «автоматически доставляет каждую крысу ко входу в лабиринт и "подбирает" их в конце без помощи экспериментатора. Объективность оценок обеспечивается применением устройства, автоматически записывающего путь крысы в лабиринте» (Tryon, 1929, p. 73). Сегодня подобная автоматизация — обычное явление. Вспомните описанное в главе 4 исследование крыс в радиальном лабиринте. В ходе эксперимента «макровыборы» и «микровыборы» сверялись по видеозаписи поведения крыс в лабиринте, а в операциональные определения этих конструктов входило условие: поведение должно легко поддаваться проверке (Brown, 1992). Кроме того, компьютеры дают возможность представлять испытуемым инструкции и стимулы и одновременно собирать данные.

Второй способ контроля за искажением, вносимым экспериментатором — это **двойной слепой метод.** Он заключается в том, что экспериментатор остается в неведении (слепой метод) о том, что ожидать от участников во время конкретного сеанса тестирования. Ни экспериментатор, ни испытуемые не знают, какое условие изучается, поэтому используется название «двойной». Двойной слепой метод применяется, когда ведущий исследователь разрабатывает эксперимент, а его коллеги (обычно аспиранты) собирают данные. Конечно, не всегда возможно применить этот метод — в качестве примера можно привести описанное в главе 3 исследование Даттона и Эрона (Dutton & Aron, 1974). Как вы помните, женщины-экспериментаторы встречали мужчин либо на опасном мосту, раскачивающемся на расстоянии 75 метров над рекой, либо на надежном мосту, проходящем в 3,5 метрах над той же рекой. Было бы довольно сложно скрыть от экспериментаторов, какое из условий эксперимента исследуется! Но во многих исследованиях применяются процедуры, при которых экспериментаторы не знают, какое условие

изучается в данный момент. При сравнении крыс, имеющих высокие и низкие способности к прохождению лабиринта, легко можно добиться того, чтобы экспериментатор не знал, какие крысы исследуются. В таком исследовании экспериментаторы, изучающие крыс, могут даже не знать, что у крыс исследуется интеллект.

Искажение, вносимое испытуемыми

От людей, участвующих в психологическом исследовании, также нельзя ожидать, что они будут действовать как машины. Они — люди, которые *знают,* что участвуют в эксперименте. В процессе получения обоснованного согласия им рассказали о сути экспериментов, но если в исследовании используется мистификация, то участники будут знать, что всех подробностей эксперимента им не сообщили. Более того, даже если мистификация не используется, испытуемые могут не поверить этому, ведь они участвуют в «психологическом эксперименте», а разве психологи не заняты постоянным «анализом» людей? Таким образом, можно сказать, что **искажение, вносимое испытуемыми,** зависит от ожиданий участников и от того, как они представляют свою роль в исследовании. Если на поведение испытуемого влияет знание того, что идет эксперимент и его участие важно для успеха исследования, возникает **хоторнский эффект,** названный так в честь известного исследования, посвященного производительности труда. Понять происхождение этого термина вам поможет вставка 6.2. Вас может удивить, что большинство историков считают, что хоторнский эффект назван неверно, а данные оригинального исследования были сильно искажены по политическим причинам.

ВСТАВКА 6.2

История — производительность труда в «Вестерн электрик»

Исследование, давшее имя хоторнскому эффекту, было проведено на заводе «Вестерн электрик» в городе Хоторне, штат Иллинойс. Оно длилось около 10 лет, с 1924 по 1933 г. Считается, что задача исследования состояла в изучении факторов, влияющих на производительность труда рабочих. Были проведены многочисленные эксперименты, из которых наиболее известна серия, названная «исследование сборки реле в тестовой комнате».

Для эксперимента со сборкой реле из группы рабочих завода были выбраны шесть женщин. Их работа заключалась в сборке реле для телефонной компании. Пять из них проводили сборку, а шестая подавала детали. Сборка реле, состоящих из 35 частей, представляла собой тяжелую, монотонную работу, отнимающую много времени. «Вестерн электрик» производил около 7 млн реле в год (Gillespie, 1988), поэтому вполне естественно, что они хотели, насколько это возможно повысить производительность труда рабочих.

Первая серия экспериментов с реле продолжалась с мая 1927 по сентябрь 1928 г. (Gillespie, 1988). За это время были изучены (а в действительности смешаны друг с другом) несколько переменных, характеризующих условия труда. Вносились изменения в расписание перерывов, общее время работы и в размер премии, выплачиваемой за определенный объем сделанной работы. Считается, что производительность труда в этой небольшой группе быстро достигла высокого уровня и оставалась неизменной, даже если условия труда ухудшались. В качестве примера всегда приводят «12-й период исследования», при котором рабочим сообщили, что рабочая неделя будет увеличена с 42 до 48 часов, а перерывы и бесплатные обеды будут отменены. Практически во всех учебниках приводятся подобные описания:

«Неважно, какие изменения вносились - увеличивали или уменьшали количество перерывов, длительность рабочего дня и т. д. - женщины производили все больше и больше телефонных реле» (Elmes, Kantiwitz, & Roediger, 1992, p. 205).

Было предположено, что рабочие сохраняют производительность труда, так как думают, что они составляют особую группу, находящуюся в центре всеобщего внимания, ведь они участвуют в эксперименте. Так возникло понятие хоторнского эффекта, состоящего в изменении поведения людей, если они знают, что их изучают. Возможно, такой эффект действительно существует, но имел ли он место в «Вестерн Электрик», мы еще не выяснили.

Внимательное изучение событий позволяет дать несколько интересных альтернативных объяснений. Во-первых, хотя обычно при описании данного исследования подчеркивается, что женщины были в восторге от работы в специальной комнате для тестирования, в действительности, двух из пяти отобранных сначала сборщиц пришлось удалить из комнаты за отказ подчиняться и низкие результаты. Про одну сборщицу сказали, что она «примкнула к большевикам» (Bramel & Friend, 1981). (Учтите, что Республика Советов в 20-х гг. только недавно появилась и «красная угроза» представляла гипотетическую опасность для Америки, что вызвало, к примеру, страх перед рабочими организациями.) Одна из двух новых женщин имела выдающиеся способности и была полна энтузиазма, так что быстро стала лидером группы. Очевидно, ее выбрали потому, что она «была самой быстрой сборщицей реле» (Gillespie, 1988, p. 122). Ее участие сильно повлияло на повышение производительности труда.

Вторая проблема, возникшая с интерпретацией данных о сборке реле, - статистическая. В«12-м периоде» производительность труда фиксировалась каждую неделю, а не каждый час, при этом рабочие трудились на 6 часов в неделю больше, чем в предыдущий тестовый период. Если бы учитывалась выработка реле в час, то было бы отмечено небольшое *снижение* продуктивности (Bramel & Friend, 1981). Также очевидно, что женщины были рассержены подобными изменениями, но боялись жаловаться, так как иначе их могли удалить из тестовой комнаты и они лишились бы премии. И наконец, возможно, что в хоторнских экспериментах рост производительности труда был вызван учетом результатов и вознаграждением высокой продуктивности (Parson, 1974).

Историки говорят, что события следует оценивать, учитывая их общий политический/экономический/институциональный контекст и исследование в Хоторне не должно быть исключением. Красивая картинка, в которой рабочие не замечают разницы в условиях труда и волнуются о том, чтобы их считали особенными, способствовала изменению взаимоотношений между людьми на производстве и привела к тому, что корпорации стали уделять большое внимание гуманному обращению с работниками, чтобы создать одну большую счастливую семью рабочих и руководителей. Но такая картинка также помогает усилить власть руководства и предотвратить повсеместное создание рабочих организаций, стремление к чему, по мнению некоторых историков (например, Bramel & Friend, 1981), и явилось истинной причиной исследований, проведенных в «Вестерн электрик».

Большинство людей, участвующих в исследованиях, желая помочь экспериментатору и способствовать получению значимых результатов, берут на себя **роль положительных испытуемых.** Этот феномен впервые был описан Орном (Orne, 1962). Конечно, бывают и исключения, но обычно участники очень хотят помочь исследователю и во имя психологии упорно выполняют повторяющиеся и скучные задания. Более того, если испытуемые сумеют узнать гипотезу исследования, они

могут пытаться подтвердить ее своим поведением. Для тех аспектов исследования, которые могут открыть гипотезу, проверяемую в ходе исследования, Ори использовал термин **наводящий признак.** Если эти признаки становятся слишком явными для испытуемых, они перестают вести себя естественно, что затрудняет интерпретацию результатов. Как определить, ведут ли себя участники естественно или они поняли гипотезу и стараются подтвердить ее своим поведением?

Орн набрал студентов для эксперимента с так называемой сенсорной депривацией (Orne & Scheibe, 1964) и продемонстрировал, как наводящие признаки могут повлиять на результаты исследования. Он предположил, что если сказать испытуемым, что их ждет участие в подобном эксперименте, то они будут ожидать стрессовой ситуации и станут реагировать соответствующим образом. Так и произошло. Участники, просидевшие 4 часа в маленькой, но уютной комнате, показывали признаки стресса, *только* если а) они подписывали документ, освобождающий экспериментатора от всякой ответственности в случае каких-либо происшествий с ними, и б) в комнате находилась «кнопка паники», которую можно было нажать, если стресс, вызванный депривацией, станет слишком сильным. Контрольным участникам не показывали документ, освобождающий исследователя от ответственности, в их комнате не было кнопки паники и у них не было создано ожидание сенсорной депривации. Они не демонстрировали признаков негативной реакции.

Возможность того, что наличие наводящих признаков повлияет на результаты исследования, учитывается при выборе меж- или внутрисубъектного плана. Участники, поведение которых исследуется при всех экспериментальных условиях, имеют большую возможность узнать гипотезу (гипотезы) исследования, поэтому наводящие признаки при внутрисубъектном плане потенциально более опасны, чем при межсубъектном. Для обоих планов наводящие признаки особенно разрушительны, если они влияют на одни экспериментальные условия и не влияют на другие, тем самым вызывая осложнение.

Кроме исполнения роли положительных испытуемых (т. е. попыток подтвердить гипотезу) испытуемые хотят, чтобы их воспринимали как компетентных, творческих, эмоционально устойчивых и т. д. людей. Они считают, что их оценивают в ходе эксперимента, а это приводит к особому состоянию, которое Розенберг (Rosenberg, 1996) назвал **боязнь оценки.** Испытуемые хотят, чтобы их оценили положительно, и поэтому начинают вести себя так, как, по их мнению, это должен делать идеальный человек. Беспокойство испытуемого о том, как он будет выглядеть, и его желание помочь экспериментатору часто способствуют одной и той же манере поведения, но иногда стремление создать благоприятное впечатление и желание быть положительным субъектом конфликтуют. Например, при исследовании готовности к оказанию помощи сообразительные участники могут догадаться, что они поставлены в такие экспериментальные условия, которые должны снижать ее вероятность. С другой стороны, альтруизм высоко ценится и даже считается героическим поведением. Желание быть положительным испытуемым (поддержать гипотезу) склоняют испытуемого к отказу от помощи, а боязнь оценки заставляет его стремиться помочь. В одном исследовании (Rosnow, Goodstadt, Suls & Gitter, 1973) было выдвинуто предположение, что при наличии выбора между подтверждением гипотезы и положительной оценкой последнее является более мощным мотивом поведения.

Контроль за искажением, вносимым испытуемыми

Для контроля за искажением, вносимым испытуемым, в первую очередь необходимо свести к минимуму наводящие признаки. Один способ сделать это — мистификация. Как мы узнали в главе 2, основная задача мистификации — побудить испытуемых вести себя более естественно. Второй подход, который обычно применяется при исследовании лекарств, заключается в использовании контрольной группы плацебо (см. главу 7). Процедура состоит в сравнении участников, подвергающихся экспериментальному воздействию (например, принимающих лекарство), и тех, кто только думает, что подвергается. Если люди в двух группах ведут себя одинаково, их поведение будет объясняться ожиданием результатов воздействия.

Второй способ обнаружить наличие наводящих признаков называется **проверка эффективности воздействия.** В исследовании с мистификацией ее можно провести во время дебрифинга, попросив участников высказать свои предположения об истинной гипотезе исследования («положительные испытуемые», однако, могут притвориться, что они ничего не подозревают). Проверку можно осуществить и в ходе самого исследования. Иногда нескольких, случайным образом выбранных участников, останавливают в разгар эксперимента и спрашивают, понятны ли инструкции, каковы их представления о происходящем и т. д. Также проверку эффективности воздействия проводят, чтобы определить: имеет ли определенная процедура эффект, который она должна иметь. Например, если процедура должна вызывать у людей тревогу (в случае, когда участникам сообщают, что они будут подвергнуты действию электрошока), выборку испытуемых могут остановить в середине эксперимента и исследовать у них уровень тревоги.

Еще один способ избежать появления наводящих признаков — провести полевое исследование. Если испытуемые не знают, что они участвуют в исследовании, маловероятно, что они станут размышлять о гипотезе и реагировать на наводящие признаки. Но конечно, как вы помните из обсуждения вопроса охраны личной жизни участников во вставке 3.1 и вопроса осведомленного согласия в главе 2, у полевых исследований немало своих проблем.

Ранее я отметил, что большинство участников исследований играют роль «положительных испытуемых», но это не всегда так. Существуют различия между теми, кто действительно добровольно участвует и интересуется экспериментом, и теми, кто случайно принял участие в исследовании и интересуется им гораздо меньше. Настоящие добровольцы обычно более сознательны и больше нуждаются в социальном одобрении (Adair, 1973). В случае, когда участие в исследовании входит в обязательную программу обучения и студентов просят выступить в роли испытуемых, различия между теми, кто участвует добровольно, и вынужденными участниками могут вызвать определенные трудности, ведь одни студенты будут настроены более энергично, чем другие. Кроме того, может возникнуть «эффект семестра»: настоящие добровольцы, которые действительно хотят участвовать в экспериментах, запишутся раньше, чем другие, не слишком заинтересованные исследованием. Если вы проведете исследование с двумя группами и группа 1 будет изучаться в первой половине семестра, а группа 2 — во второй, то нельзя будет определить, чем вызваны обнаруженные различия, независимой переменной или разницей между настоящими добровольцами, записавшимися первыми, и испытуемыми поневоле,

как можно дольше откладывающими свое участие. Попробуйте найти решение этой проблемы. Если вам пришла в голову идея «блоковой рандомизации» и вы подумали, что «с ее помощью можно равномерно распределить условия исследования по всему семестру», то поздравляю, вы не зря читали эту главу.

В завершение главы прочитайте вставку 6.3, касающуюся этических обязанностей участников психологических исследований. Список этих обязанностей основан на предположении, что исследование представляет собой результат совместных усилий экспериментаторов и участников. Мы уже знаем, что экспериментаторы должны следовать положениям этического кодекса *APA.* Вставка 6.3 расскажет вам о том, что участники также имеют определенные обязанности.

ВСТАВКА 6.3

Этика — участники исследования также имеют обязанности

Этический кодекс *APA* описывает обязанности исследователей перед участниками своих экспериментов. Участники вправе ожидать выполнения обязательств, а в случае их невыполнения должна быть предусмотрена четкая процедура для выражения протеста. Но как обстоят дела со стороной участников? Каковы их обязательства?

В статье Джеймса Корна, напечатанной в журнале *Teaching of Psychology* (1988), приводятся не только основные права, которые имеют студенты колледжей, участвующие в исследованиях, но и их обязанности, в том числе:

1) ответственно относиться к расписанию - не пропускать встречи с исследователями и приходить вовремя;
2) содействовать проведению исследования, действовать профессионально и честно, прикладывая максимальные усилия;
3) внимательно слушать экспериментатора во время процедуры получения обоснованного согласия и при инструктаже и задавать вопросы, если остается неясной суть дела;
4) уважать требование исследователя избегать обсуждений эксперимента до окончания сбора данных;
5) быть активным при проведении дебрифинга, помогая исследователю понять изучаемое явление.

Этот список основывается на предположении, что исследование должно представлять собой результат совместных усилий экспериментаторов и участников. Можно одобрить идею Корна о том, что участники должны играть более активную роль, чтобы сделать сотрудничество исследователя и испытуемых более полным. Однако эта активность должна сопровождаться внимательным изучением того, что участники думают об исследовании. Экспериментатор, который просто «управляет испытуемым» и собирает данные, проходит мимо ценной информации.

В двух предыдущих главах вы познакомились с основными чертами экспериментальных исследований и некоторыми проблемами контроля, которые необходимо изучить всем, кто собирается заниматься психологическими исследования-

ми. Мы завершили предварительную работу, необходимую для знакомства с различными видами экспериментальных планов, которые применяются для выявления действия независимых переменных, и можем приступить к более подробному изучению этих планов.

Резюме

Межсубъектные планы

При использовании межсубъектного плана испытуемые изучаются только при одном из экспериментальных условий, а следовательно, каждое условие требует участия новой группы испытуемых. Обычно такой план используется, если изучаются субъектные переменные (например, пол) или если выполнение заданий при одном условии изменяет испытуемых и делает невозможным их участие при изучении других условий. Главная проблема исследований с межсубъектным планом — создание групп, эквивалентных друг другу по всем показателям, кроме независимой переменной.

Проблема создания эквивалентных групп

Наилучший способ создания эквивалентных групп при использовании межсубъектного плана — это случайное распределение. Случайное распределение равномерно размещает непредвиденные осложнители по разным группам, а следовательно, предотвращает их разрушительное воздействие. Вероятность того, что случайное распределение будет эффективным, возрастает с увеличением количества участников в каждой группе. Если количество участников очень мало, а определенный фактор (например, интеллект) в высокой степени коррелирует с зависимой переменной и при этом этот фактор несложно оценить до начала эксперимента, эквивалентные группы могут быть созданы с помощью процедуры уравнивания.

Внутрисубъектные планы

Если каждый участник изучается при всех экспериментальных условиях, в исследовании используется внутрисубъектный план или план с повторяемыми измерениями. В таких исследованиях участие в экспериментах с одним условием может повлиять на поведение участников при других условиях, а следовательно, вызвать эффект последовательности или эффект порядка. Отсутствие контроля за такой ситуацией может привести к смешению разнородных результатов исследования. Разновидностями эффекта последовательности являются эффект прогрессии (постепенно накапливающийся, как в случае усталости) и эффект передачи (одна последовательность условий может давать результаты, отличающиеся от полученных при другой последовательности).

Проблема контроля за эффектом последовательности

Эффект последовательности можно контролировать с помощью различных процедур позиционного уравнивания, обеспечивающих использование разных последовательностей условий при их изучении. Если испытуемые участвуют в экспери-

ментах с разными условиями по одному разу, используется полное (изучаются все возможные последовательности) или частичное (выборка, состоящая из разных последовательностей, или латинский квадрат) позиционное уравнивание. Если участие осуществляется при каждом из условий более одного раза, может использоваться обратное позиционное уравнивание или блоковая рандомизация. При наличии эффекта передачи возникает асимметричный перенос, снижающий эффективность процедуры позиционного уравнивания.

Проблемы контроля в исследованиях развития

В психологии развития основной независимой переменной является возраст, представляющий собой субъективную переменную. При проведении межсубъектных исследований возраст изучается методом поперечных срезов. Данный метод весьма эффективен, но при его использовании может возникнуть эффект когорты — особый вид проблемы неэквивалентных групп. Если возраст является внутрисубъектной переменной, то план называется лонгитюдным. Его основной проблемой является истощение. Эти два подхода можно скомбинировать, если каждые несколько лет выбирать новую когорту и исследовать каждую когорту длительное время.

Проблемы искажения

Результаты психологического исследования могут искажаться ожиданиями экспериментатора. Эффект ожидания может привести к тому, что экспериментатор начнет при разных условиях по-разному воздействовать на испытуемых, что сделает невозможным верную интерпретацию результатов. Этот эффект .может быть снижен автоматизацией процедуры исследования и использованием двойного слепого метода. Может также возникнуть искажение, вносимое испытуемыми. Если наводящие признаки подскажут участникам истинную цель экспериментов, то они могут попытаться подтвердить гипотезу исследования, а осознавая свое участие в эксперименте, испытуемые могут изменить манеру поведения. Наводящие признаки обычно держатся под контролем путем варьирования степени мистификации, а искажение, вносимое испытуемыми, оценивается с помощью проверки эффективности воздействия.

Задания для повторения

Выбор ответа

1. Основное преимущество межсубъектного плана перед внутрисубъектным заключается в том, что межсубъектный план:
   1) требует участия меньшего количества испытуемых;
   2) по определению не позволяет появиться проблеме эквивалентных групп;
   3) снижает уровень дисперсии между различными условиями;
   4) по определению не позволяет появиться эффекту последовательности.
2. Какое из следующих утверждений о блоковой рандомизации верно?
   1) ее используют, чтобы добиться уравнивания;

2) это другое название латинского квадрата;
3) при межсубъектном плане ее можно использовать, чтобы обеспечить одинаковое количество участников в каждой группе;
4) при внутрисубъектном плане так называется процедура полного позиционного уравнивания.

3. В каком случае для создания эквивалентных групп лучше использовать уравнивание, чем случайное распределение?
   1) всегда при большом количестве испытуемых;
   2) при наличии потенциального осложнителя;
   3) если известно, что внешняя переменная, которую возможно измерить, коррелирует с зависимой переменной;
   4) при наличии очевидных наводящих признаков.
4. При возникновении эффекта асимметричного переноса:
   1) позиционное уравнивание может не справиться с устранением эффекта последовательности;
   2) необходимо использовать частичное, а не полное позиционное уравнивание;
   3) необходимо использовать полное позиционное уравнивание;
   4) необходимо использовать уравнивание, а не случайное распределение.
5. Эффект ожидания экспериментатора:
   1) был обнаружен в исследованиях с участием людей, а в случае исследований с участием животных его обнаружить не удалось;
   2) может быть снижен путем автоматизации процедуры исследования;
   3) не возникает, если участники не знают гипотезы, которая проверяется в ходе исследования;
   4) никогда не воспроизводился после исследования Розенталя, поэтому, возможно, такой проблемы не существует.

Короткие эссе

1. При каких обстоятельствах лучше использовать межсубъектный план, чем внутрисубъектный?
2. При каких обстоятельствах лучше использовать внутрисубъектный план, чем межсубъектный?
3. Чем случайный выбор отличается от случайного распределения и какова задача последнего?
4. В какой ситуации для создания эквивалентных групп скорее всего будет использовано уравнивание?
5. Назовите отличия эффекта прогрессии от эффекта передачи и объясните, почему позиционное уравнивание более эффективно в случае первого, чем в случае последнего.

6. При тестировании вкуса Джоан попросили попробовать и оценить четыре вида белого сухого вина: *A, B, C и D.B* каком порядке их будут пробовать при использовании а) обратного позиционного уравнивания и б) блоковой рандомизации? Сколько последовательностей будет создано, если исследователь использует полное позиционное уравнивание?
7. Каковы характерные особенности латинского квадрата и в каких ситуациях его чаще всего используют?
8. Какие особые проблемы контроля возникают в психологии развития при использовании а) метода поперечных срезов и б) лонгитюдных исследований?
9. Приведите пример исследования, иллюстрирующего искажение, вносимое экспериментатором. Как можно контролировать такое искажение?
10. Что такое наводящие признаки и как их можно контролировать?

Упражнения

Упражнение 6.1. Межсубъектные или внутрисубъектные планы?

Представьте себе исследования, в которых проверяются приведенные ниже гипотезы. В частности, укажите, какой, по вашему мнению, должна быть независимая переменная, меж- или внутрисубъектной, и будет ли разумным проведение подобных исследований. Объясните свой ответ.

1. Нейрофизиолог выдвигает гипотезу о том, что повреждение основной зрительной зоны мозга у взрослых животных невосстановимо.
2. Физиолог, занимающийся изучением восприятия, предполагает, что близкие оттенки серого цвета легче различать при естественном освещении, чем при искусственном.
3. Клинический физиолог считает, что наилучший способ лечения фобии состоит в том, чтобы показывать человеку вызывающий страх объект и не позволять уйти от воздействия до тех пор, пока он не поймет, что объект безопасен.
4. Ученый, занимающийся психологией развития, предполагает, что культура, к которой принадлежит человек, влияет на его моральное развитие.
5. Социальный психолог думает, что люди подходят к решению проблем более творчески, находясь в коллективе, чем в одиночку.
6. Когнитивный психолог выдвигает гипотезу о том, что изучение вербальной информации по частям приводит к лучшему запоминанию, чем изучение всего материала целиком.
7. Клинический специалист предполагает, что люди с синдромом навязчивых состояний легче поддаются гипнозу, чем люди с различными фобиями.

Упражнение 6.2. Построение правильного латинского квадрата

Ученый, исследующий память, хочет сравнить долговременную память на примере нескольких списков слов и определить ее зависимость от того, изучает человек в первый раз четыре списка или восемь. Помогите ученому спланировать исследование, построив два латинских квадрата с помощью процедуры, описанной в табл. 6.3.

Упражнение 6.3. Использование блоковой рандомизации

Экспериментатор хочет проверить гипотезу о том, что социальное положение жертвы влияет на приговор преступника, совершившего изнасилование. Участникам раздали описание преступления и попросили определить наказание для признанного виновным преступника. Разным группам сказали, что жертве был 21 и она была проституткой, матерью двоих детей, студенткой, не имевшей сексуального опыта, женщиной в возрасте 21 года (контрольная группа). В каждой группе было по 5 участников. С помощью блоковой рандомизации распределите участников по четырем группам и составьте список, в котором укажите, при каком из условий будет исследоваться каждый из 20 участников.

Упражнение 6.4. Случайное распределение и уравнивание

Исследователь изучает сравнительную эффективность двух программ потери веса. Участников распределяют на две экспериментальные группы и одну контрольную группу листа ожидания. Чтобы в начале исследования люди в одной из групп не оказались заметно тяжелее, чем в другой, было решено уравнять все три группы по начальному весу. Ниже приведен вес в фунтах всех 15 участников.

156	167	183	170	145
143	152	145	181	162
175	159	169	174	161

Сначала с помощью процедуры уравнивания сформируйте эквивалентные группы. Затем воспользуйтесь случайным распределением и еще раз распределите участников по группам (примите каждый столбец за «блок» и с помощью блоковой рандомизации поместите равное количество участников в каждую группу). Сравните эти два подхода, подсчитав по два значения среднего арифметического для каждой из групп (по одному для каждого подхода). Сравните ваши результаты с результатами, полученными другими учащимися вашей группы. Какой вывод можно сделать об уравнивании и случайном распределении?

ГЛАВА 7

Экспериментальный план I: однофакторные планы

Обзор задач главы

Главы 5 и 6 образуют сцену для этой и следующих глав. В главе 5 мы познакомились с экспериментальным методом, получили представление о независимых, внешних, и зависимых переменных, изучили проблему осложнений и обсудили факторы, связанные с валидностью психологических экспериментов. В главе 6 мы сравнили межсубъектные и внутрисубъектные планы, изучили основные способы контроля за ними (например, случайное распределение и позиционное уравнивание), а также рассмотрели проблему искажений, возникающих в психологических исследованиях. Сцена готова, и очень скоро на ней будет разыграно представление этой и следующих глав. Они расскажут вам о различных экспериментальных планах, на основании которых в психологии строятся экспериментальные исследования. В данной главе рассматриваются планы с одной независимой переменной, принимающей одно или несколько значений. Если используется более одной независимой переменной, то план называется факторным — такой план рассматривается в главе 8. Изучив данную главу, вы:

- сможете определить четыре вида однофакторного плана: планы с независимыми, уравненными и неэквивалентными группами, а также план с повторяемыми измерениями;
- научитесь понимать, когда нужно применять проверку по ^-критерию Стьюдента для независимых и зависимых групп;
- научитесь понимать, когда нужно использовать различные виды контрольных групп: группы плацебо и листа ожидания, а также сцепленную контрольную группу;
- поймете этические аспекты использования контрольных групп;
- сможете назвать две главные причины использования более чем двух значений для независимой переменной;
- сумеете решать, когда нужно использовать гистограмму, а когда линейный график;
- поймете, почему при анализе данных однофакторного многоуровневого эксперимента не стоит применять многократную проверку по критерию Стьюдента, а следует использовать метод *ANOVA*.

В главе 3 при обсуждении научной креативности в качестве примера была рассказана история происхождения исследования с лабиринтом. Исследование Смолла, в котором использовалась модель лабиринта Хэмптон Корт, было первым из целой лавины экспериментов с лабиринтами, проведенных в течение первых двух десятилетий XX в. В основном в ранних исследованиях ученые стремились выяснить, какое из чувств крысы наиболее важно для процесса обучения. Вы, вероятно, помните, как в главе ? (вставка 2.3) рассказывалось о проблемах, возникших у Джона Уотсона с борцами против вивисекции из-за серии исследований, в ходе которых он хирургическим путем последовательно лишал крыс одного чувства за другим. Уотсон обнаружил, что способность к изучению лабиринта не снижается, даже если крыса лишается большинства чувств, и сделал вывод, что для запоминания лабиринта крысы пользуются мускульным, или кинестетическим, чувством: они запоминают, что нужно сделать определенное количество шагов, а затем повернуть направо и т. д. Для проверки этой идеи он вместе со своим коллегой из Чикагского университета Харви Карром провел простое, но изящное исследование (Carr & Watson, 1908). После того как группа крыс изучила сложный лабиринт, Карр и Уотсон удалили центральную часть лабиринта, сделав один из его участков короче. Они предположили, что крысы, натренированные проходить более длинный лабиринт, будут буквально натыкаться на стены укороченного лабиринта. И действительно, в описании поведения одной из крыс Карр и Уотсон отметили, что она «на полном ходу натыкалась на стены. Сильно шаталась и не вернулась к нормальному поведению, пока не прошла (еще) 9 футов» (р. 39). Вторая группа крыс обучалась на более коротком лабиринте, а затем переходила к более длинному. Эти крысы вели себя сходным образом, нередко слишком рано поворачивали и натыкались на стены, по всей видимости, ожидая найти поворот. Спустя многие годы после ухода из академической науки, Джон Уотсон назвал это исследование одним из важнейших в своей жизни. Последующие эксперименты с использованием лабиринтов поставили под сомнение кинестетическую гипотезу, но для нас в этой истории важно то, что хорошее исследование не всегда нуждается в чрезвычайно сложном экспериментальном плане. В некоторых случаях двух групп вполне достаточно.

Один фактор — два уровня

Как показано на рис. 7.1, изображающем дерево принятия решений, есть четыре вида плана с одной независимой переменной, которая принимает два значения (т. е. для которой имеется два уровня). К каждому из этих видов приводит последовательность решений, принятых в отношении независимой переменной. Во-первых, эта переменная может быть внутрисубъектной или межсубъектной. Если ее изучают как межсубъектную, она может быть управляемой или субъектной. При управляемой независимой переменной план будет называться **планом с независимыми группами,** если для создания эквивалентных групп используется случайное распределение, и **планом с уравненными группами,** если для этого используется уравнивание, а затем случайное распределение. Как вы узнали в главе 6, чтобы при-

нять решение об использовании уравнивания, необходимо учитывать размер выборки и особенно внимательно относиться к внешним переменным, коррелирующим с зависимой переменной. Если изучается субъектная переменная, группы формируются из разных категорий людей (например, женщины/мужчины, интроверты/экстраверты, либералы/консерваторы). Экспериментальный план при этом иногда называют *«ex post facto»*, так как группы формируются после установления наличия у испытуемых определенных характеристик. Такой план также называют планом со «стихийно возникшими группами», или **планом с неэквивалентными группами** (я буду использовать именно этот термин). Поскольку при таком плане группы строятся из разных типов людей, исследователи нередко пытаются по возможности снизить неэквивалентность, уравнивая группы по различным факторам. Например, в исследовании с неэквивалентными группами, в котором сравниваются женщины и мужчины, можно сделать так, чтобы члены обеих групп были одного возраста и принадлежали к одному социоэкономическому классу.

Последний вид однофакторного плана — это **план с повторяемыми измерениями. Он** используется, если независимая переменная является внутрисубъектной и каждый участник исследуется при каждом значении независимой переменной (т. е. измерения повторяются несколько раз). Важнейшие особенности каждого из четырех основных видов экспериментального плана представлены в табл. 7.1. Далее мы приступим к рассмотрению конкретных примеров.

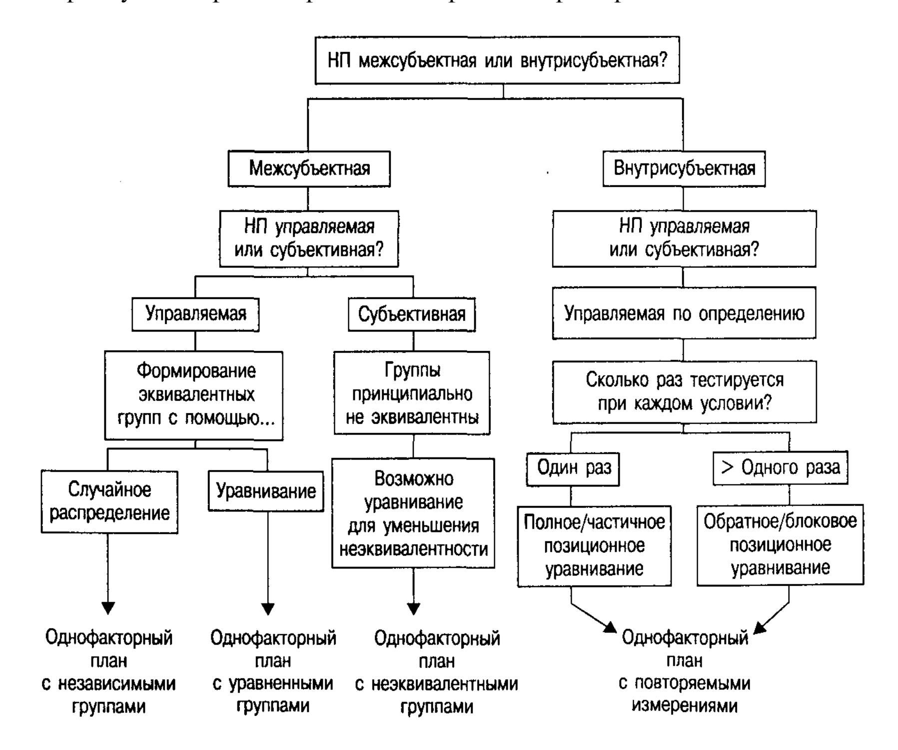


Рис. 7.1. Древо решений - однофакторный план

Таблица 7,1

Особенности четырех однофакторных планов

Вид плана	Минимальное количество значений, принимаемых независимой переменной	Межсубъектная или внутрисубъектная независимая переменная	Вид независимой переменной	Способ создания эквивалентных групп
С независимыми группами	2	Межсубъектная	Управляемая	Случайное распределение
С уравненными группами	2	Межсубъектная	Управляемая	Уравнивание
С неэквивалентными группами	2	Межсубъектная	Субъектная	Уравнивание может снизить неэквивалентность
С повторяемыми измерениями	2	Внутрисубъектная	Управляемая	Отсутствует

Межсубъектные однофакторные планы

Однофакторные двухуровневые исследования проводятся не так часто, как вы можете решить. Большинство исследователей предпочитают использовать более сложные планы, позволяющие получить более подробные и интересные результаты. Кроме того, мало кто из редакторов журналов достаточно высоко оценит такое простое исследование. Но несмотря на это, в простоте есть особая прелесть, а что может быть проще сравнения двух условий? Ниже приводятся три примера таких исследований.

Пример 6. Независимые группы

Примером эксперимента, проведенного по однофакторному плану с независимыми группами, является хорошо известное исследование Блэкмора и Купера (Blackemore & Cooper, 1970). Ученые заинтересовались вопросом влияния опыта на развитие зрительной системы. Двухнедельных котят случайным образом распределили по двум значениям независимой переменной, которую можно обозначить как «зрительное окружение». Котят растили в условиях с преобладанием либо вертикальных либо горизонтальных полос. Я думаю, вы понимаете, почему исследование должно было быть межсубъектным, а не внутрисубъектным — было бы бессмысленно растить котят в «вертикальном» окружении, а *затем* в «горизонтальном». В подобных исследованиях испытуемые, принявшие участие в эксперименте с одним значением независимой переменной, уже по сути «использованы» и полученный опыт не дает им возможности «начать сначала» с другим условием.

На рис. 7.2 показан выполненный Блэкмором и Купером набросок устройства для создания «вертикального» окружения. Кот стоит на площадке из оргстекла, с которой видно, как полосы уходят вверх и вниз. Широкий воротник на шее жи-

вотного заставляет его зрительно фокусироваться на стенах камеры. В течение нескольких месяцев котов помещали в «вертикальный» или «горизонтальный» миры на 5 часов в день; остальное время они находились в темноте.

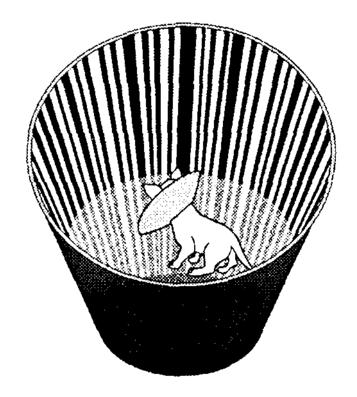


Рис. 7.2. Устройство для эксперимента Блэкмора и Купера (Blakemore & Cooper, 1970)

В конце эксперимента Блэкмор и Купер исследовали поведение животных и измерили у них активность нейронов зрительной зоны коры головного мозга. В целом коты быстро оправились от депривации. «Через 10 часов пребывания в нормальных зрительных условиях они... смогли с легкостью прыгать со стула на пол» (Blackemore & Cooper, 1970, p. 477). Однако коты, выросшие в «вертикальном» окружении, не очень хорошо воспринимали явления, происходящие в горизонтальной плоскости, а вертикальные стимулы вызывали проблемы у животных, привыкших к «горизонтальным» условиям:

> Разница стала особенно заметна, когда двух котят, из которых один имел «горизонтальный», а второй — «вертикальный» опыт, одновременно исследовали с помощью длинной белой или черной рейки. Если рейку держали вертикально и трясли из стороны в сторону, один из котят подбегал и начинал играть с ней. Если рейку держали горизонтально, это привлекало другого котенка, а первый оставался безразличным.
>
> Blackmore & Cooper, 1970, p. 478

Очевидно, что опыт первых месяцев жизни сильно влияет на развитие мозга.

Пример 7. Уравненные группы

В старых фильмах иногда показывают, как героя держат взаперти, пытают и лишают сна на 2 или 3 дня, чтобы получить от него определенные сведения. Может ли депривация сна повлиять на ответы, полученные от человека в ходе допроса? Такой эмпирический вопрос поставил перед собой Блэгров (Blagrove, 1996) при проведении интересного исследования с уравненными группами. Выражаясь более точно, он хотел узнать, как подействуют на людей, лишенных сна, вводящие в заблуждение вопросы. Исследователь набрал студентов колледжа для проведения трех различных экспериментов, в каждом из которых должны были участвовать по две группы — одни участники подвергались депривации сна, а другие — нет. Лишенные сна сту-

денты оставались в лаборатории и бодрствовали по 21 часу подряд в ходе первых двух экспериментов и 43 часа в третьем. Постоянное наблюдение «20 смен ассистентов» (р. 50) гарантировало, что испытуемые не спят. Студенты, которых не подвергали депривации, могли спать дома. Переменной уравнивания была «обычная продолжительность сна, сообщаемая испытуемыми» (р. 50). Блэгров хотел уравнять обычную продолжительность сна, чтобы иметь возможность «контролировать у групп с одинаковой продолжительностью сна различия личностных и связанных с продолжительностью сна характеристик» (р. 50). Средняя продолжительность сна составила 8,4 и 8,5 в первом исследовании, 8,3 и 8,1 во втором и 8,4 и 8,1 в третьем.

Все участники исследования выполнили стандартизованный тест на внушаемость: они прослушали рассказ, а затем ответили на наводящие вопросы (т. е . такие вопросы, на которые нельзя ответить прямо, используя информацию из рассказа). Ответив на вопросы, они получили негативную оценку выполнения задания, после чего их попросили еще раз ответить на те же вопросы, чтобы увидеть, изменят ли они какие-либо из ответов. В целом, на лишенных сна испытуемых вопросы действовали сильнее и они чаще меняли свои ответы. Особенно ярко это проявилось в третьем исследовании, в котором депривация сна продолжалась 43 часа. Процедура уравнивания помогла создать группы, сходные по продолжительности сна, что позволило объяснить различия между группами именно депривацией.

Пример 8. Неэквивалентные группы

Вероятно, под влиянием мегалонгитюдного исследования Термана по изучению одаренных детей (см. вставку 6.1), Кнеппер с соавторами (Knepper, Obrzut & Copeland, 1983) провели заслуживающее внимания исследование, в котором попытались пролить свет на личностные особенности одаренных детей. Исследователи поставили перед собой вопрос, успешнее ли одаренные дети помимо познавательных задач решают социальные и эмоциональные проблемы в сравнении с обычными детьми. Их эксперимент отлично иллюстрирует план с неэквивалентными группами. Независимой субъективной переменной была степень одаренности. Сравнивались два ее значения: одаренный (операционально определенное как *IQ* = 130 и выше) и средний (*IQ* между 90 и 110). Среднее арифметическое значение *IQ* составило 136,9 и 102,9 соответственно. Никакой особой процедуры уравнивания использовано не было, но возраст контролировался — в исследовании принимали участие только шестиклассники. Использовался тест *Means-Ends Problem Solving Test,* оценивающий качество решений межличностных (социальных) и внутриличностных (эмоциональных) проблем. Одаренные дети действительно показали более высокие результаты, чем обычные, — это открытие совпадает с выводом Термана о том, что одаренные дети не просто «умные», а имеют также и социальные навыки.

Необходимо сделать одно важное замечание. Вспомните, как в главе 5 рассказывалось о том, что вывод о наличии причинно-следственной связи нельзя сделать, если используются субъектные переменные. Поэтому неверно говорить, что одаренность каким-либо образом вызывает увеличение способности к решению социальных и эмоциональных проблем. Можно лишь сказать, что у одаренных и неодаренных детей способности к решению подобных проблем различаются.

Внутрисубъектные однофакторные планы

Как вы уже знаете, внутрисубъектые планы: а) требуют меньшего количества участников, б) более чувствительны к небольшим различиям в значениях среднего арифметического и в) обычно предполагают использование позиционного уравнивания для устранения проблем последовательности. В случае внутрисубъектного плана с одной независимой переменной, принимающей два значения, можно использовать два вида позиционного уравнивания. Если испытуемые участвуют в экспериментах один раз при каждом условии, применяется полное позиционное уравнивание. Тогда половина участников выполняет задания при условии А, а затем В, а вторая половина — при условии В, а затем А. Если участники исследуются более одного раза при каждом условии, можно применить обратное позиционное уравнивание (АВВА). Именно этот подход использовал Дж. Ридли Струп в первом из трех проведенных им исследований, результаты которых были опубликованы в 1935 г. В списке «10 лучших классических исследований» его исследование заняло бы далеко не последнее место. Чтобы получить о нем более подробную информацию (и побольше узнать о том, что такое свастика), прочитайте вставку 7.1.

ВСТАВКА 7.1

Классические исследования — наиболее часто воспроизводимое психологическое открытие

Именно обратное позиционное уравнивание Дж. Ридли Струп использовал в своей работе, впервые опубликованной им в 1935 г. Это исследование настолько известно, что феномен, который был в нем продемонстрирован, сегодня называется «эффектом Струпа». В сопроводительной статье к повторной публикации данной работы Колин Мак-Леод назвал эффект Струпа «золотым стандартом» измерения внимания и начал свое эссе такими словами: «практически невозможно найти человека, занимающегося когнитивной психологией и при этом не получившего хотя бы краткого знакомства с эффектом Струпа. Эти слова, вероятно, относятся ко всем, кто прошел стандартный вводный курс по психологии - в нем тест Струпа демонстрируется практически обязательно». (MacLeod, 1992, р. 12).

Далее Мак-Леод указал, что эффект Струпа в психологии - это одно из наиболее часто воспроизводимых и упоминаемых открытий. В чем же состоит это открытие?

В исследовании обобщались три эксперимента, проведенные Струпом в ходе работы над докторской диссертацией. Мы остановимся на первых двух, поскольку они иллюстрируют внутрисубъектный план с одной независимой переменной, принимающей два значения, и поскольку в них использовалось обратное позиционное уравнивание. В первом эксперименте 14 женщин и 56 мужчин выполняли два задания, каждое из которых включало чтение названий цветов. Первое условие Струп назвал ЧНЦч («чтение названий цветов, напечатанных черным») (Stroop, 1992, р. 12). Участники должны были прочитать 100 названий цветов (например, ЗЕЛЕНЫЙ), напечатанных черными чернилами, так быстро и точно, как только возможно. Второе условие Струп обозначил как ЧНЦр («чтение названий цветов, когда цвет чернил и название различны») (р. 16). В этом случае 100 названий цветов были напечатаны цветными чернилами, но при этом цвет чернил не соответствовал названию цвета (например, слово ЗЕЛЕНЫЙ было напечатано красным). Задание состояло в чтении слова (т. е. правильный ответ - «зеленый»).

Как хороший исследователь, Струп знал о проблеме последовательности и поэтому для ее устранения использовал обратное позиционное уравнивание (АВВА). Разделив каждый из списков стимулов на отрезки, состоящие из 50 слов, Струп выдал половине участников последовательность ЧНЦч-ЧНЦр-ЧНЦр-ЧНЦч, а второй половине - последовательность ЧНЦр-ЧНЦч-ЧНЦч-ЧНЦр. Таким образом, каждый из участников должен был в целом прочитать 200 названий цветов.

В первом эксперименте Струп не обнаружил *никаких различий* между выполнением заданий с условиями ЧНЦч и ЧНЦр. Среднее время, требующееся для прочтения 100 слов каждого вида, составило 41,0 и 43,3 секунды соответственно. Таким образом, на чтении цветных названий в условии ЧНЦр не сказалось то, что слова были напечатаны контрастным цветом.

Но во втором эксперименте Струп обнаружил огромные различия, и это впоследствии сделало его имя знаменитым. Используя тот же общий план эксперимента, он изменил задание - на этот раз необходимо было называть цвета, а не читать их названия. При одном условии - НЦ («тест на называние цвета»), - участники называли цвет образцов квадратной формы. При втором, ключевом условии - НЦСр («тест на называние цвета слова, когда цвет чернил и слово различны») - участникам показывали тот же материал, что и при условии ЧНЦр из эксперимента 1, но в этот раз, вместо того чтобы читать название цвета, они должны были называть цвет, которым напечатано слово. Если слово ЗЕЛЕНЫЙ было напечатано красным, правильный ответ был «красный», а не «зеленый». В 1935 г. участники встретились с теми же трудностями, что переживают испытуемые сегодня. Поскольку чтение представляет собой в высшей степени глубоко усвоенный, автоматический процесс, он мешает называнию цвета, что приводит к ошибкам и замедлению выполнения задания. Струп обнаружил, что среднее время называния цветов составило 63,3 секунды для условия НЦ и (поразительно) 110,3 секунды для условия НЦСр. Из четырех результатов, представленных Струпом в виде таблиц, я построил гистограмму, показанную на рис. 7.3. Хорошо видно, что эффект Струпа весьма силен.

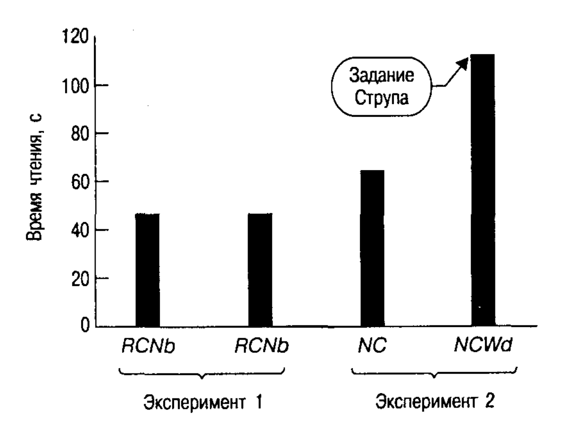


Рис. 7.3. Данные первых двух экспериментов из исследования Струпа

Как я уже отмечал, в действительности в ходе работы над диссертацией Струп провел три эксперимента. Третий из них показывает, что при наличии тренировки испытуемые могут улучшить выполнение задания НЦСр (классическое задание Струпа). Интересная особенность последнего эксперимента заключалась в том, что цветные квадратные образцы из НЦ-теста Струп заменил цветными образцами в форме свастики, что «позволило печатать НЦ-тес,т используя опенки, более близкие к использованным в НЦСр-тесте» (Stroop, 1992, p. 18). Свастика - это древний религиозный символ, который получается, если загнуть перекладины традиционного греческого креста (+). Так вышло, что исследование Струпа было опубликовано в том же году (1935), когда свастика стала официальным символом нацистской Германии.

Другой способ позиционного уравнивания в случае исследования с двумя условиями, при каждом из которых участники исследуются несколько раз, — простое чередование условий (ABAB...). Этот подход рассматривается в следующем примере.

Пример 9. Повторяемые измерения

Для проверки нескольких предположений, сделанных на основании выдвинутой Джеймсом Гибсоном теории восприятия, Ли и Аронсон (Lee & Aronson, 1974) провели исследование, посвященное восприятию движения и чувству равновесия (о Джеймсе Гибсоне, муже Элеанор Гибсон, упоминалось в главе 1). В частности, они интересовались тем, как мы сохраняем равновесие в движущейся среде. Исследователи помещали детей в возрасте от 13 до 16 месяцев в аппарат, изображенный на рис. 7.4. Когда ребенок смотрел на дальнюю стену, экспериментатор двигал стены и потолок вперед или назад.

Была выдвинута гипотеза, что движение комнаты вперед (рис. 7.5, а) создает «оптический эффект движения», аналогичный возникающему при движении головы ребенка назад (рис. 7.5, б). Это вызовет у ребенка компенсаторный наклон вперед. Если предположение верно, то движение комнаты вперед должно привести к тому, что ребенок наклонится или даже упадет вперед (рис. 7.5, б). Подобное предположение было сделано для движения комнаты в противоположную сторону.

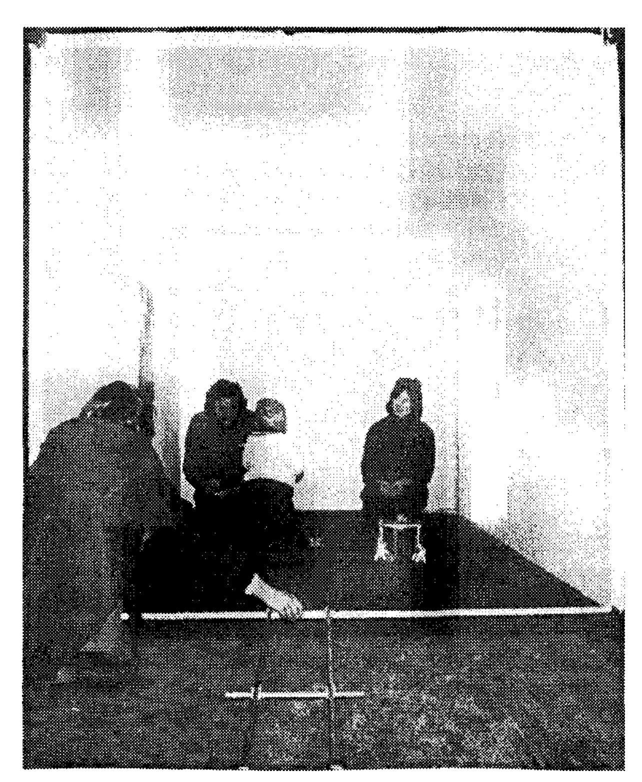


Рис. 7.4. Аппарат, использованный в эксперименте с движущейся комнатой, проведенном Ли и Аронсоном в 1974 г.

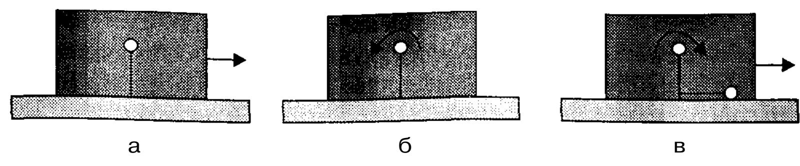


Рис. 7.5. Предсказанное влияние движения комнаты вперед в эксперименте Ли и Андерсона 1974 г.

В противоположность исследованию с кошками, выращенными в «вертикальном» или «горизонтальном» мире, в данном случае не было препятствий для исследования одних и тех же детей при обоих экспериментальных условиях — движении комнаты назад и вперед. Поэтому использовался внутрисубъектный однофакторный план с повторяемыми измерениями. Независимой переменной было направление движения комнаты (вперед или назад), а наклон тела или падение ребенка оценивалось в качестве зависимой переменной. Каждый испытуемый выполнил по двадцать попыток, при этом направление движения комнаты чередовалось от попытки к попытке. У одних детей последовательность начиналась с движения комнаты вперед, а у других — с движения назад. Было набрано семь участников в возрасте от 13 до 16 месяцев, но трое из них испытали сильный стресс и для них эксперимент был прекращен. Реакции оставшихся четырех участников записывались тремя наблюдателями (почему понадобилось более одного наблюдателя?). Потеря равновесия в предсказанном направлении наблюдалась в 82% всех попыток. Наблюдатели выделяли три вида потери равновесия: качание (26% попыток), шатание (23%) и падение (33%).

Один недостаток процедуры позиционного уравнивания, состоящей в простом чередовании условий, (а) и (в), заключается в том, что испытуемые могут легко предсказать, какое условие будет использоваться следующим. Однако Ли и Аронсон справедливо решили, что при подобном возрасте участников эта проблема не повлияет на результаты исследования. Другая причина использования такой разновидности позиционного уравнивания была практической: поскольку ребенок оставался в комнате в течение всего сеанса, если комнату двигали в одну сторону, то при следующей попытке ее необходимо было двигать обратно.

Анализ однофакторных двухуровневых планов

Чтобы выяснить, являются ли различия, обнаруженные между результатами двух условий при двухуровневом плане, значимыми или они вызваны случайностью, требуется провести статистический анализ вывода. Если в обработке результатов эксперимента используется интервальная шкала измерений или шкала отношений, для проведения анализа чаще всего применяется одна из двух разновидностей проверки по критерию Стьюдента. Данная процедура упоминалась ближе к концу главы 4, а более подробно описывается в приложении *С*. При использовании номинальной или порядковой шкал необходимо использовать другие подходы.

Есть две разновидности проверки по критерию Стьюдента. Первая из них называется **проверкой для независимых групп** и, как ясно из названия, она используется, если две группы участников совершенно независимы друг от друга. Это возможно: а) когда участники исследования случайным образом распределены по

группам или б) если изучаемая переменная является субъективной (например, если участники делятся на женщин и мужчин). Если независимая переменная представляет собой внутрисубъектный фактор или две различные группы людей формируются таким образом, что между ними устанавливаются определенные отношения (например, испытуемые в группе Л уравнены по интеллекту с участниками группы *В),* то используется проверка по критерию Стьюдента для зависимых групп (иногда ее называют проверкой по критерию Стьюдента для коррелированных групп). Четырем рассмотренным однофакторным планам соответствуют следующие варианты проверки по критерию Стьюдента:

- проверка по критерию Стьюдента для независимых групп:
  - план с независимыми группами;
  - план с неэквивалентными группами;
- проверка по критерию Стьюдента для зависимых групп:
  - план с уравненными группами;
  - план с повторяемыми изменениями.

По сути, проверка по критерию Стьюдента заключается в нахождении разницы между двумя значениями среднего арифметического и выяснении (с определенной вероятностью), не превышает ли эта разница значения, получаемого при воздействии только случайных факторов. Если она больше этого значения и можно исключить потенциальные осложнители, то исследователь может с высокой вероятностью заключить, что различия действительно существуют, опубликовать исследование и, возможно, получить повышение по службе. В приложении С приведена пошаговая инструкция для обоих видов проверки по критерию Стьюдента и для анализа силы эффекта.

Планы с контрольными группами

Понятие об экспериментальных и контрольных группах, а также о различиях между ними было дано в главе 5. Экспериментальные группы подвергаются определенному воздействию, а контрольные — не подвергаются. В случае плана с повторяемыми измерениями, не подразумевающего наличия разных групп, данное рассуждение переносятся на условия исследования, среди которых выделяются экспериментальные и контрольные. Кроме обычной контрольной группы, не подвергающейся воздействию, следует отметить еще три особых вида контрольных групп: контрольные группы плацебо, листа ожидания и сцепленные группы.

Контрольные группы плацебо

«Плацебо» — это вещество, имеющее определенный эффект, но при этом фармакологически неактивное. Иногда пациентам становится лучше, когда они принимают плацебо, но думают, что получают лекарство X Это происходит потому, что люди верят в то, что лекарство им поможет. В исследованиях члены **контрольной группы плацебо** думают, что подвергаются определенному воздействию, но в действительности этого не происходит. Далее вы поймете, почему может потребоваться использовать такой подход. Предположим, вы хотите выяснить, замедляет ли

алкоголь реакцию человека. Если вы используете обычную экспериментальную группу, которую попросите принять алкоголь, и вторую, сохраняющую трезвость, а затем определите у испытуемых скорость реакции, то может оказаться, что реакция у первой группы будет замедленной. Можно ли будет заключить, что алкоголь снижает скорость реакции? Нет, ведь участники могут придерживаться распространенного мнения, что алкоголь замедляет реакцию, и это знание может оказать на них некоторое влияние. Чтобы избавиться от такого искажения, вам потребуется использовать еще одну группу и дать ей напиток, кажущийся алкогольным (который на вкус невозможно отличить от алкогольного), но таковым не являющийся. Такая группа является контрольной группой плацебо. Следует ли устранять обычную контрольную группу (вообще не принимающую алкоголь)? Не обязательно, ведь эти испытуемые могут дать образец скорости реакции. Если вы используете все три группы и получите следующие значения средней скорости реакции:

экспериментальная группа — 0,32 с;
контрольная группа плацебо — 0,22 с;
обычная контрольная группа — 0,16 с,

то сможете сделать вывод, что ожидание людей по поводу действия алкоголя несколько замедляет скорость реакции (с 0,16 до 0,22 с), но и сам алкоголь также оказывает влияние (снижение скорости реакции с 0,22 до 0,32 с).

Примером эксперимента с двумя контрольными группами — обычной и плацебо — является исследование влияния заражения паразитами на различные познавательные и моторные навыки у детей с Ямайки (Sternberg, Powell, McGrane & Grantham-McGregor, 1997). Ученики четвертых и пятых классов с легкой степенью заражения были случайным образом распределены (после уравнивания по полу) на две группы: получающую лекарство и плацебо. Дети из третьей (неэквивалентной) группы не были заражены. Лекарство быстро устранило инфекцию, и по окончании эксперимента его дали также и детям из группы плацебо. Но последствия заражения остались. В сравнении с незараженной контрольной группой, не получающей лечения, дети из двух других групп плохо справились с заданиями на проверку когнитивных способностей. Следовательно, лекарство имело важный медицинский эффект, но мало способствовало усилению когнитивных способностей, снизившихся из-за болезни. Поскольку дети жили в условиях, где паразитические инфекции весьма распространены, Стернберг и др. порекомендовали включить в лечение программу коррекции когнитивных способностей.

Контрольные группы листа ожидания

Контрольные группы листа ожидания нередко используются в исследованиях по оценке эффективности программ (глава 10) или выявлению действенности психотерапии. Такой план используется, когда члены экспериментальной группы участвуют в программе из-за того, что испытывают определенные проблемы, решению которых эта программа должна помочь. Например, в исследовании Миллера и Дипилато (Miller & DiPilato, 1983) оценивалась эффективность двух видов терапии (релаксации и десенситизации) для людей, страдающих ночными кошмарами. Исследователи решили использовать контрольную группу, не подвергающуюся воз-

действию, но для того, чтобы испытуемые во всех трех группах были в целом эквивалентны друг другу, члены контрольной группы также должны были страдать от ночных кошмаров. Участникам, попавшим в группу листа ожидания, сказали, что им обязательно помогут, и по окончании исследования им обеспечили лечение, эквивалентное тому, которому были подвергнуты члены экспериментальных групп.

То, что членам группы листа ожидания обеспечивалось лечение, было весьма полезно для здоровья испытуемых, но также это оказало определенное давление на исследователей, которые были вынуждены использовать данную контрольную процедуру только для программ, имеющих относительно небольшую продолжительность. Могут возразить, что использование контрольной группы листа ожидания неэтично, так как люди сначала не получают никакой пользы от программы. Этот вопрос может стать особенно сложным, если в ходе исследования оценивается программа, влияющая на жизнь человека. Более подробно этот вопрос рассматривается во вставке 7.2, защищающей использование контрольных групп в исследованиях.

ВСТАВКА 7.2

Этика — кто войдет в контрольную группу?

При проведении исследования памяти, в котором экспериментальную группу просят формировать визуальные образы, тогда как контрольная группа просто заучивает список слов, вопрос о том, кто должен быть в контрольной группе, не создает этической дилеммы. Однако если эксперимент разрабатывается для оценки некоторой программы или терапевтического воздействия, которое может принести заметную пользу людям или даже продлить их жизнь, то этот вопрос решить нелегко. Например, в известном исследовании влияния личного контроля на здоровье (Langer & Rodin, 1976) нескольким людям, живущим в доме для престарелых, позволили самим планировать свой день, тогда как члены контрольной группы пользовались (в основном) распорядком, установленным для них персоналом. Когда авторы через 18 месяцев вернулись и провели повторное исследование, они обнаружили, что в среднем в сравнении с контрольной группой члены первой группы были более психически и физически здоровы и большее количество из них были живы (Rodin & Langer, 1977). Если бы вы узнали, что один из ваших родственников попал в контрольную группу, вы бы обеспокоились?

Точно так же возникли споры с распределением участников по группам в исследовании с больными раком пациентами (Adler, 1992). В ходе исследования изучалось влияние групп поддержки на психологическое и физическое состояние женщин, страдающих раком груди. Было обнаружено, что женщины в группах поддержки быстрее выздоравливали и даже жили дольше, чем другие (т. е. распределенные в контрольную группу). Некоторые исследователи утверждали, что результаты отражают не пользу, приносимую группами поддержки, а вред, нанесенный членам контрольной группы, которые могли почувствовать себя обделенными или отвергнутыми. Такие чувства могли вызвать стресс, а как известно, стресс неблагоприятно влияет на иммунную систему, что приводит к разнообразным проблемам со здоровьем. Так есть ли хоть доля правды в рис. 7.6? Может ли участие в контрольной группе убить кого-либо?

Защитники использования контрольных групп при оценке программ имеют три сильных аргумента. Во-первых, они отмечают, что легко оценивать прошедшие события и говорить уже после свершившегося факта, что «программа, настолько эффективная, как эта, должна быть доступна для всех». Проблема в том, что *до* «свершившегося факта» далеко не так очевидно то, что программа будет эффективной. Единственный способ узнать это - провести исследование. Например, до начала исследования Лэнгера и Родена в доме для престарелых мож-

но было легко предположить, что участники эксперимента испытают ненужный стресс от того, что должны сами о себе заботиться, и быстрее умрут. Точно так же защитники исследования рака указывают, что в начале исследования довольно мало женщин имели предпочтения по поводу распределения по группам, а несколько из них просто не захотели присоединиться к группе поддержки (Adler, 1992). Следовательно, неверно, что члены контрольной группы чувствовали себя отвергнутыми или обделенными.

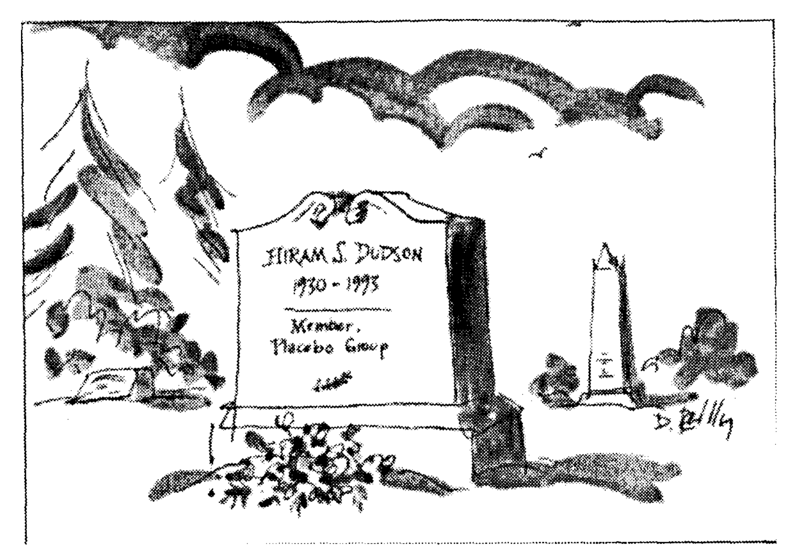


Рис. 7.6. Возможные последствия попадания в контрольную группу?
Надпись на рисунке: *Хайрем С. Дадсон, 1930-1993, член группы плацебо*

Во-вторых, экспериментаторы указывают, что в исследованиях по оценке новых программ или способов терапии редко сравниваются новый способ воздействия и отсутствие воздействия - обычно проводится сравнение нового метода воздействия со старым. Таким образом, членам контрольной группы доступны методы лечения, которые нормально применяются, и, соответственно они получают обычную помощь. Более того, если исследование показывает наличие положительного эффекта от экспериментального воздействия, члены контрольной группы обычно получают возможность испытать его на себе.

В-третьих, лечение обходится дорого и стоит потратить деньги на лучшее из возможного. Однако нельзя определить без подробно разработанного исследования эффективности программы, какой метод лучше. Программы, эмпирически подтвердившие свою эффективность, служат на благо всем людям и нередко могут сохранить или продлить им жизнь.

Помните, как в главе 1 обсуждался вопрос псевдонауки и приводился в связи с этим пример «бессознательно воспринимаемых аудиозаписей»? Как мы узнали, было проведено обширное исследование, показавшее, что положительное влияние этих записей возникает в результате ожидания их эффективности. Ожидания нередко выявляются путем использования контрольных групп плацебо, но мы рассмотрим исследование, в котором использовались одновременно и группы плацебо, и группы листа ожидания. Это исследование позволило по-новому интерпретировать возможности подобных записей.

Пример 10. Использование контрольных групп плацебо и листа ожидания в одном исследовании

Один из самых распространенных вариантов использования «бессознательно воспринимаемых аудиозаписей» — это использование их с целью снижения веса. Эта область предоставляет очень хорошие возможности для продажи подобной продукции. Американцы, к примеру, чтобы сбросить вес, испытывают на себе всевозможные приемы, от диеты с ограниченным потреблением жиров до хирургического вмешательства. Особенно привлекательными для людей оказываются методы, требующие минимальных усилий, а именно это и является отличительной чертой «бессознательно воспринимаемых аудиозаписей» — просто вставьте кассету и очень скоро ваше бессознательное начнет руководить вашим поведением и неизбежно приведет к потере веса. Для изучения эффективности таких записей в деле снижения веса Мерикл и Скейнс (Merikle & Skanes, 1992) провели исследование, в котором использовали и контрольные группы плацебо, и группы листа ожидания. Дав объявление в газеты, они набрали сорок семь женщин и случайным образом распределили их в три группы. Члены экспериментальной группы *(N*= 15) получили коммерческую «аудиозапись бессознательного действия», которая, по мнению производителей, должна была помочь слушателям сбросить вес. Члены контрольной группы плацебо *(N=15)* думали, что получили аудиозаписи, специально разработанные для снижения веса, но в действительности им выдали записи для снижения страха лечения зубов (исследователи обладали чувством юмора). Для обычных слушателей эти две записи были неразличимы. Членам контрольной группы листа ожидания *(N*= 17) сказали, «что в исследовании уже принимает участие максимальное количество испытуемых и что... их поставят на лист ожидания» (р. 774). Участников из экспериментальной группы и группы плацебо попросили прослушивать записи 1-3 часа в день; при этом членов всех трех групп на протяжении пяти недель еженедельно взвешивали. Каковы были результаты? Потеря веса у членов экспериментальной группы была весьма скромной, но *столько же* сбросили члены контрольной группы плацебо. Это обычный результат для подобного типа исследований, показывающий, что «бессознательно воспринимаемые аудиозаписи» сами по себе не имеют никакого эффекта. Однако интересно, что контрольная группа листа ожидания также потеряла в весе, и примерно столько же, сколько остальные две группы. На основании этого Мерикл и Скейнс заключили, что «бессознательно воспринимаемые записи» дают предполагаемый результат не просто из-за наличия эффекта плацебо. Если бы все дело было в эффекте плацебо, то члены группы плацебо, думая, что на их разум оказано воздействие, сбросили бы вес, а вес испытуемых из группы листа ожидания, которые еще не слушали записи, не изменился бы. Полученные результаты привели авторов к выводу, что воздействие записей заключалось в концентрации внимания людей на существующей проблеме, в данном случае на снижении веса. Испытуемые во всех трех группах «могли потерять вес просто потому, что участие в исследовании увеличило вероятность того, что в ходе эксперимента они станут внимательно относиться к вопросам, связанным со снижением веса» (р. 776). В данном исследовании группа листа ожидания позволила оценить силу эффекта плацебо и получить альтерна-

тивное объяснение популярности «бессознательно воспринимаемых записей». Также, хотя авторы и не отметили этого, результаты исследования выглядят подозрительно похожими на хоторнский эффект, о котором вы узнали в главе 6.

Следует отметить еще один момент, касающийся этого исследования: второй автор, Хизер Скейнс, проводил эксперименты, а первый, Филипп Мерикл, наклеивал этикетки на кассеты. Таким образом, он был единственным, кто знал, какие участники получают записи для потери веса (экспериментальная группа), а какие — записи для снижения страха лечения зубов (группа плацебо). Тем самым исследователи применили двойной слепой метод.

Эквивалентные контрольные группы

Третий вид контрольных групп — **эквивалентные контрольные группы.** Их используют, если члены экспериментальной группы по какой-либо причине участвуют в эксперименте разное количество времени или включены в разные события, происходящие в ходе исследования. Каждый участник из контрольной группы уравнен, или «сцеплен», с одним из членов экспериментальной группы, таким образом для групп в целом затраченное участниками время, а также разновидности событий, с которыми они встречаются, поддерживаются постоянными. Приведенный ниже пример поможет вам лучше понять идею эквивалентных контрольных групп.

Пример 11. Эквивалентные контрольные группы и стресс

Хороший пример эквивалентной контрольной группы можно найти в исследовании Вейсса (Weiss, 1968), посвященном изучению взаимосвязей между управлением стрессорами и здоровьем. К хвостам крыс в ходе эксперимента в случайном порядке посылали слабые электрические разряды. Крысы могли выключить ток (т. е. управлять им) или избежать его воздействия, вращая лапами небольшое колесо (рис. 7.7). Контрольная группа крыс не подвергалась шоку. Крысы, находящиеся в центральной части аппарата, принадлежали к сцепленной контрольной группе. Каждая из них была сцеплена (имела парой) с крысой из экспериментальной группы и получала такое же количество электрических разрядов, но не могла управлять ими. Каждая попытка начиналась с сигнала, который слышали все три животных. У крысы из экспериментальной группы было 10 секунд, чтобы повернуть колесо. Если она не делала этого, то к ее хвосту посылался электрический разряд. При этом сцепленная крыса также получала разряд (вне зависимости от того, что она делала в это время). Таким образом, крысы в экспериментальной и контрольной группах получали абсолютно одинаковое количество раздражителей, вызывающих отрицательную реакцию (т. е. одинаковое количество электрических разрядов), но отличались по возможности контроля за ситуацией. Вейсс заключил, что возможность управлять поступлением тока помогла крысам избежать некоторых разрушительных для здоровья последствий стресса. У крыс из сцепленной группы вероятность развития язвы и потери веса была выше, чем у крыс из экспериментальной и контрольной групп, на которых подобный опыт не оказал почти никакого влияния.

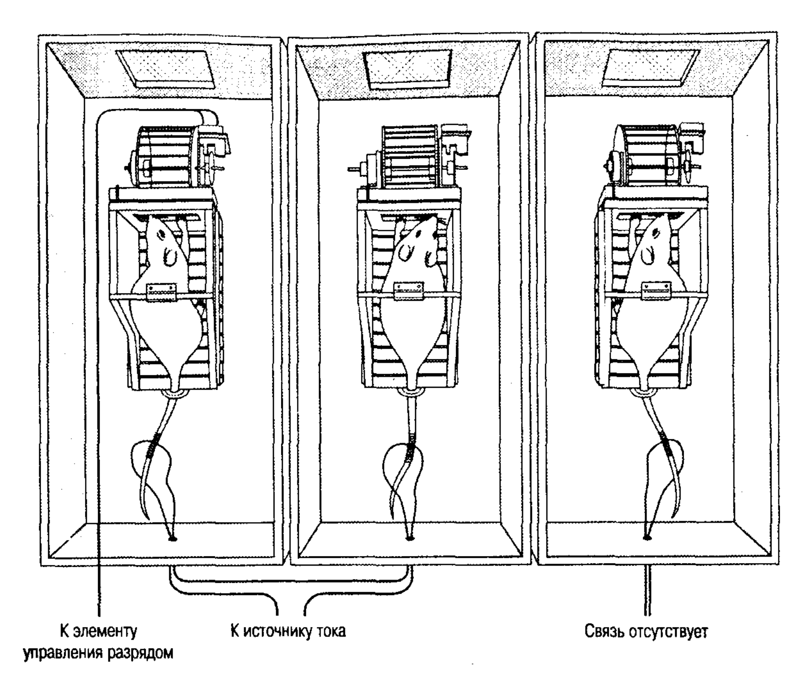


Рис. 7.7. Данные первых двух экспериментов из исследования Вейсса

Вы уже знакомы с экспериментальным планом, использованным Вейссом, — его использовала исследовательская группа Селигмана при изучении выученной беспомощности. Если вы еще раз обратитесь к обсуждению теорий, приведенному в главе 3, то обнаружите описание процедуры с эквивалентными контрольными группами, использованной в триадном плане Селигмана. Другой пример эквивалентной контрольной группы дает исследование Брэди, посвященное изучению язвы у «исполнительных» обезьян, описанное во вставке 5.3. По сути, при разработке своего исследования Вейсс скорректировал методологические недостатки (т. е. проблему отбора испытуемых), обнаруженные в исследовании Брэди (Weiss, 1977).

Один фактор — более двух уровней

Если в экспериментах используется одна независимая переменная, ситуация, когда изучаются только два ее значения, является скорее исключением, чем правилом. В большинстве однофакторных исследований используется три или более значений независимой переменной, поэтому такие планы часто называют **однофакторными многоуровневыми планами.** Так же как и двухуровневые, многоуровневые планы могут быть меж- и внутрисубъектными, а также планами с независимыми, уравненными или неэквивалентными группами, или планами с повторяемыми измерениями.

Межсубъектные многоуровневые планы

Явное преимущество многоуровневых планов заключается в том, что они позволяют исследовать нелинейные эффекты. Рассмотрим простой пример межсубьектного плана. Предположим, вы интересуетесь влиянием различных доз кофеина на скорость реакции. Вы разрабатываете эксперимент, в котором сравниваете два уровня дозировки (1 и 3 мг), получаете результаты, представленные на рис. 7.8, и делаете вывод, что кофеин как стимулятор ускоряет реакцию. При увеличении дозировки скорость реакции нарастает по прямой (т. е. линейно).

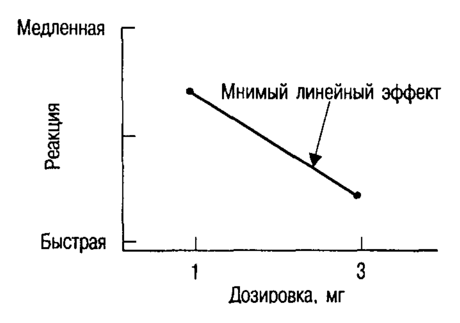


Рис. 7.8. Гипотетическое влияние кофеина на скорость реакции - два уровня

Далее предположим, что другой исследователь провел такое же исследование, но использовал многоуровневый план, включающий четыре дозы кофеина (1, 2, 3 и 4 мг) — пример повторения (1 и 3 мг) и дополнения (2 и 4 мг) вашего эксперимента. Такое исследование может дать результаты, представленные на рис. 7.9.

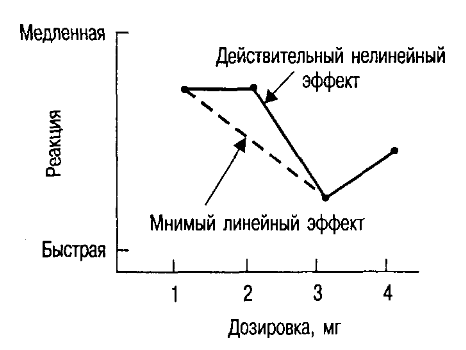


Рис. 7.9. Гипотетическое влияние кофеина на скорость реакции - четыре уровня

Таким образом, в точности воспроизведены ваши результаты для условий с дозировкой в 1 и 3 мг, но в целом результаты для четырех условий ставят ваши выводы под сомнение. Вместо вывода о том, что кофеин увеличивает скорость реакции, в данном случае можно заключить, что а) увеличение дозы кофеина повышает скорость реакции, но лишь по достижении уровня в 2 мг и б) кофеин увеличивает скорость реакции только до определенного момента — после 3 мг кофеин снижает ее. Это значит, что результаты больше не показывают линейную зависимость и являются нелинейными. В целом, преимущество многоуровневых планов состоит

в том, что они более информативны и зачастую дают более интересные результаты, чем двухуровневые.

Один из самых известных в психологии графиков иллюстрирует нелинейный эффект. Он показывает, как время, прошедшее с момента запоминания, влияет на объем забытого материала, и объясняет, почему через два дня после экзамена очень часто говорят «Я ничего не помню!» Эта кривая впервые была получена в новаторском исследовании памяти, проведенном Германом Эббингхаузом (см. вставку 7.3).

ВСТАВКА 7.3

История — нелинейные результаты - кривая забывания Эббингхауза

Немецкий психолог XIX в. Герман Эббингхауз (1850-1909) известен своим новаторским исследованием памяти и забывания. В те времена, когда психология переживала свое детство и руководств о том, «как проводить психологические исследования», еще не существовало, Эббингхауз провел серию исследований, ставших стандартом точности и методологической строгости. Цель исследования заключалась в изучении образования ассоциаций в человеческом сознании, и первой задачей было найти материал, свободный от ассоциаций. Способ решения этой задачи является одним из лучших примеров креативности в психологии: ученый образовал последовательности, состоящие из согласных, гласных и снова согласных. Такие последовательности, обозначаемые как CVC (consonants, vowels, consonants), более известны как бессмысленные слоги. Эббингхауз создал около 2300 таких слогов. В течение нескольких лет, демонстрируя потрясающую настойчивость, а возможно, и полное отсутствие личной жизни, Эббингхауз запоминал, а затем старался вспомнить последовательности CVC. Да, действительно, он был единственным испытуемым. Он систематически изменял такие факторы, как количество слогов в списке, количество экспериментальных попыток на один список и время, разделяющее попытки друг от друга. Эббингхауз опубликовал результаты своего исследования в небольшой монографии, названной «Память: исследование по экспериментальной психологии» (Memory: A Contribution to Experimental Psychology, 1885/1964).

Одно из самых известных своих исследований, в ходе которого были получены нелинейные результаты, Эббингхауз посвятил изучению процесса забывания с течением времени. Был поставлен следующий эмпирический вопрос: какой объем из запомненного материала сохраняется спустя различные отрезки времени? Процедура состояла в том, чтобы запомнить восемь списков по 13 бессмысленных слогов в каждом, подождать некоторое время, а затем попытаться выучить их еще раз. Временные интервалы составили 20 минут, 1 час, 9 часов, 1 день, 2 дня, 6 дней и 31 день. Эббингхауз зафиксировал общее время первоначального и повторного заучивания восьми списков. Их разность составила величину «экономии», которую ученый перевел в проценты, разделив на время первичного запоминания. Таким образом, если начальное запоминание заняло 20 минут, а повторное - 5 минут, то было сэкономлено 15 минут или 75% (15 / 20 x 100), от времени первичного запоминания.

Результаты, полученные Эббингхаузом, показаны на рис. 7.10. Вероятно, вы также сможете найти их в главе, посвященной памяти, учебника по общей психологи. Несложно заметить, что процент запомненных слогов снижался со временем, но это снижение не носило равномерного или линейного характера. Очевидно, что эффект был нелинейным. Сначала забывание было очень сильным, но затем оно замедлялось. Таким образом, через 20 минут было сэкономлено только около 60% (58,2) от времени первичного запоминания. Другая часть кривой показывает, что между интервалами в неделю (сэкономлено 25,4%) и месяц (сэкономлено 21,1%) не было больших отличий.

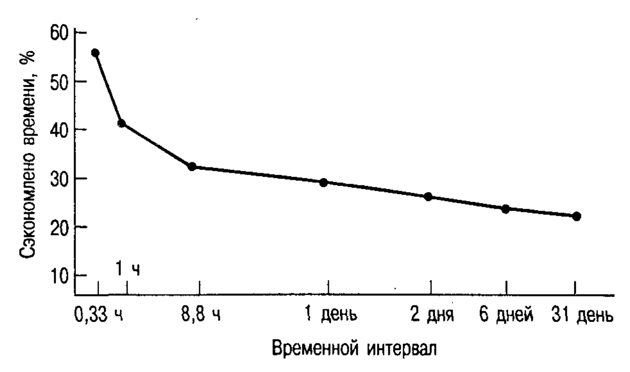


Рис. 7.10. Кривая забывания Эббингхауза - нелинейные результаты

Есть несколько интересных методологических особенностей исследования Эббингхауза. Например, чтобы сделать постоянным время показа слогов, ученый установил метроном на 150 ударов в минуту и читал каждый слог CVC в течение ровно одного удара. Также он старался изучать списки слогов в одних и тех же условиях, в одно время суток и не использовать никаких особых техник для запоминания, кроме простого повторения. Кроме того, он работал, только когда чувствовал достаточную мотивацию и мог «концентрировать внимание на утомительном задании и его цели» (Ebbinghaus, 1885/1964, p. 25).

Кроме выявления нелинейного эффекта однофакторные многоуровневые планы также используются для проверки альтернативных гипотез. Именно этот подход приводился в пример в главе 3 при обсуждении положительных сторон фальсификации. Превосходный пример этого являет собой исследование Брэнсфорда и Джонсона (Bransford & Jonson, 1972).

Пример 12. Многоуровневый план с независимыми группами

Ученые, посвятившие себя когнитивной психологии и занимающиеся вопросами понимания новой для нас информации, показали, что понять новую идею легче, если она окружена определенным контекстом. Например, главу из книги будет легче понять, если вы сначала прочитаете обзор и задачи этой главы. Я надеюсь, вы уже обнаружили это при работе с данным учебником. Исследование Брэнсфорда и Джонсона показывает влияние контекста на понимание. В ходе исследования участников просили прочитать и понять следующий текст. Попробуйте и вы:

Если шарики лопнут, звук не сможет быть услышан, поскольку это произойдет слишком далеко от нужного этажа. Закрытое окно также не даст возможности распространиться звуку, так как большинство зданий обладают весьма неплохой звукоизоляцией. Поскольку вся процедура зависит от равномерного потока электричества, разрыв провода также может вызвать проблемы. Конечно, человек может кричать, но человеческий голос недостаточно силен, чтобы разнестись так далеко. Дополнительная проблема состоит в том, что у инструмента может порваться струна. Тогда сообщение останется без аккомпанемента. Очевидно, что чем меньше будет расстояние, тем лучше. Тем самым будет меньше потенциальных проблем. При личном контакте возможностей потерпеть неудачу будет гораздо менъпе(Bransford & Jonson, 1972, p. 392).

Я думаю, вашей реакцией на этот пассаж будет что-то вроде «Хм?» Именно так реагировало большинство участников данного исследования. Однако Брэнсфорд и Джонсон обнаружили, что пониманию можно помочь, использовав определенный контекст.

Психологи разработали однофакторное исследование с независимыми группами, в котором независимая переменная принимала пять значений. Испытуемые, случайным образом распределенные в контрольную группу, выполняли задание, аналогичное только что выполненному вами: они читали текст и старались запомнить как можно больше идей из 14, содержащихся в тексте. В результате в среднем они запомнили 3,6 идей, что совсем не впечатляет. Чтобы проверить, сможет ли простое повторение улучшить запоминание, вторую группу попросили прочитать рассказ дважды. Это не помогло — испытуемые запомнили 3,8 идей. Третьей группе предварительно показали рисунок (рис. 7.11, а), а затем попросили прочитать и пересказать текст. Участники из этой группы запомнили 8,0 идей из 14. Очевидно, что рисунок явился общим контекстом, помогшим участникам понять рассказ. Но обязательно ли *сначала* рассмотреть рисунок, а затем прочитать текст? Да, обязательно. При четвертом условии эксперимента участники читали рассказ, *затем* рассматривали рисунок, а потом пересказывали прочитанный текст. Они запомнили 3,6 идей — ровно столько, сколько запомнили члены контрольной группы. Пятой группе предложили частичный контекст. Прежде чем читать рассказ, они рассмотрели рис. 7.11,6, содержащий все отдельные части рисунка 7.11, а, но иначе расположенные. Испытуемые из этой группы запомнили в среднем 4,0 идей. Полученные результаты можно изобразить графически, как показано на рис. 7.12.

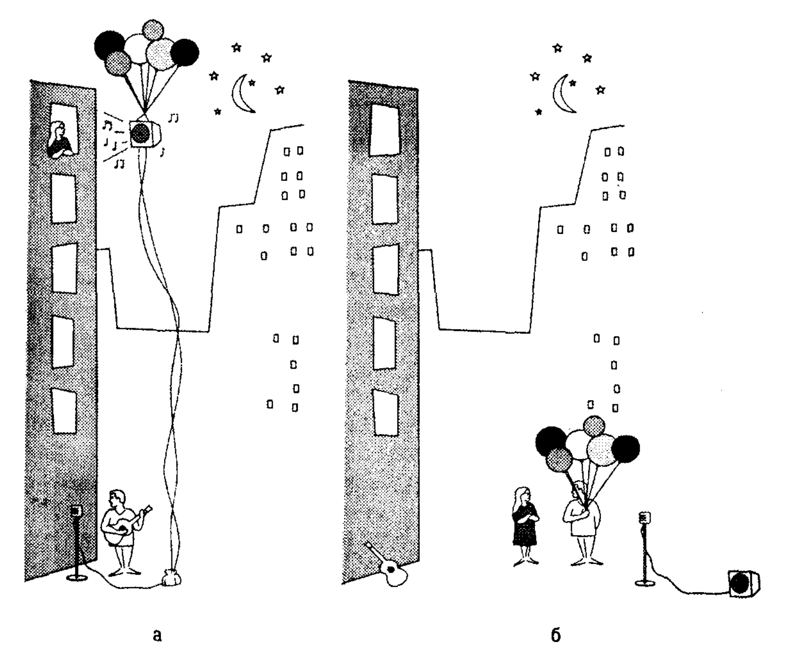


Рис. 7.11. Рисунок, обеспечивающий: а) контекст и б) частичный контекст для исследования Брэнсфорда и Джонсона (Bransford & Jonson, 1972)

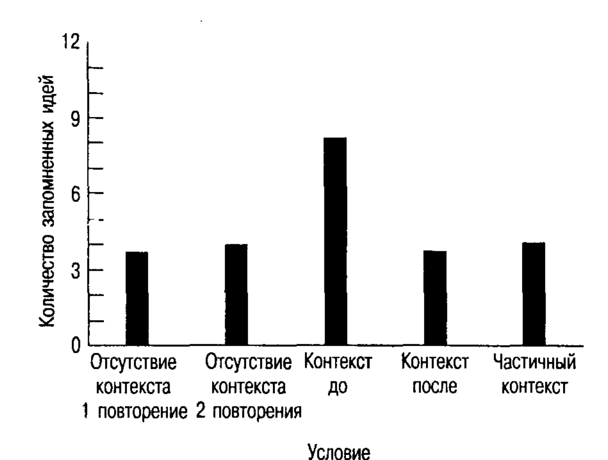


Рис. 7.12. Данные по пяти условиям исследования Брэнсфорда и Джонсона, представленные в виде гистограммы

Аналогичное исследование с двумя группами: «Отсутствие контекста — одно повторение» и «Контекст до», было бы весьма интересным и показало бы улучшение понимания при введении контекста в форме рисунка. Но использование всех четырех условий делает исследование *действительно* интересным и позволяет отбросить (т. е. фальсифицировать) некоторые альтернативные факторы, которые иначе будут расценены как улучшающие запоминание. Таким образом, контекст улучшает понимание, но *только* если он введен предварительно. Так как представление контекста после прочтения текста не улучшает запоминания, можно сделать вывод, что контекст делает это, облегчая первичную обработку информации, а не ее последующее восстановление в памяти. Можно также отбросить предположение, что простое повторение улучшает запоминание, ведь удваивание количества повторений текста не способствовало запоминанию. Кроме того, недостаточно просто показать отдельные части рассказа (как в случае с частичным контекстом) — они должны быть расположены в определенном порядке и отражать суть запоминаемого материала.

Внутрисубъектные многоуровневые планы

Тогда как в случае однофакторного двухуровневого плана с повторяемыми измерениями можно применить лишь отдельные варианты позиционного уравнивания, выход за пределы двух уровней дает возможность использовать все его виды. Если каждое условие изучается один раз для каждого испытуемого, доступны и полное, и частичное позиционное уравнивание. Если каждое условие исследуется несколько раз для каждого испытуемого, можно использовать обратное позиционное уравнивание или блоковую рандомизацию. В нижеследующем примере каждое условие изучалось только один раз и для позиционного уравнивания использовался латинский квадрат.

Пример 13. Многоуровневые планы с повторяемыми измерениями

Может ли прослушивание музыки Моцарта сделать вас умнее? Некоторые считают, что может, — подобный феномен был назван «эффект Моцарта». Несмотря на отсутствие фактов, поддерживающих эту идею, родителей пытаются убедить, что

музыка Моцарта способна дать их детям оружие для IQ-битвы. Существует даже веб-сайт, посвященный эффекту Моцарта (www.mozarteffect.com), где можно заказать различные записи и книги. В истинно псевдонаучной манере сайт доказывает действенность этого явления, приводя отзывы покупателей и единичные свидетельства. В описании одной из записей, продающейся на сайте (концерта для скрипки), утверждается, что прослушивание концерта «увеличит вербальные и эмоциональные способности, улучшит концентрацию внимания, память и навыки интуитивного и пространственного мышления», а также что «высокая частота звуков скрипки тренирует слух, а следовательно, стимуляция мозга уравновешивается прекрасной гармонией» (что бы это ни значило). Каковы же основания для таких необычайных утверждений? Поиски приводят нас к небольшому исследованию, опубликованному в 1993 г. в журнале *Nature*. В нем было показано, что прослушивание произведений Моцарта по 10 минут в день производит кратковременное (т. е. эффект длится недолго) усиление способности к пространственному мышлению у студентов колледжа (Rauscher, Shaw, & Key, 1993). Занимающиеся когнитивной психологией ученые скептически отнеслись к информации даже о таком кратковременном эффекте и неоднократно пытались повторить это исследование. Все попытки были неудачными. Одно из таких исследований провели Стил, Болл и Ранк (Steele, Ball & Runk, 1977).

В своем исследовании Стил и его коллеги использовали три условия: прослушивание произведений Моцарта в течение 10 минут, прослушивание успокаивающих звуков естественных природных явлений (например, мягкий шум дождя) в течение 10 минут и отсутствие прослушивания каких-либо звуков — в течение 10 минут участники в сидели в тишине и старались расслабиться. Все 36 участников исследовались при каждом условии, а следовательно, план был межсубъектным многоуровневым. Хотя легко можно было провести полное позиционное уравнивание (шесть различных последовательностей условий, шесть участников, случайным образом распределенные в каждой из шести последовательностей), авторы решили использовать латинский квадрат размером 3x3, подразумевающий участие 12 испытуемых, случайно распределенных в каждом из рядов квадрата. Чтобы предотвратить искажение, которое могло возникнуть, если бы участники знали, что изучается эффект Моцарта, им «сказали, что эксперимент посвящен влиянию релаксации на запоминание» (Steel et al., 1997, p. 1181). Задание на проверку памяти заключалось в воспроизведении в обратном порядке набора чисел. Если стимул был «6-8-3-1-7», то правильный ответ должен был быть «7-1-3-8-6». В ходе каждой попытки испытуемые слушали Моцарта, мягкие звуки дождя или сидели в тишине, а затем выполняли три задания на запоминание чисел. Каждое задание состояло из 9 чисел, представленных в случайном порядке. Таким образом, участники могли набрать от 0 до 27 очков.

Результаты исследования являются статистически значимыми, но они отнюдь не способствуют продажам записей Моцарта. Среднее количество правильно запомненных чисел было практически одинаковым для всех трех условий: 18,53 для записей Моцарта, 18,50 для записи шума дождя и 18,72 для контрольного условия. Однако наблюдался значительный эффект тренировки. Вне зависимости от порядка следования условий участники продемонстрировали улучшение результатов от первого набора чисел для запоминания к третьему (средние оценки для них соста-

вили соответственно 15,64,19,14 и 20,97). Так нужно ли родителям ставить записи произведений Моцарта своим детям? Конечно, нужно, ведь это прекрасная музыка. Сделает ли она их умнее? Нет, но зато они насладятся классической музыкой, что само по себе хорошо.

Представление данных

При составлении отчетов о результатах исследований необходимо решить, каким образом представить данные. Есть три возможности. Во-первых, их можно представить в повествовательной форме — такой подход хорош, если экспериментальное исследование включает два или три значения независимой переменной (как, например, в исследовании эффекта Моцарта), но будет весьма скучным при увеличении количества данных. Возможно, вы уже заметили это, когда читали о результатах исследования Брэнсфорда и Джонсона, содержащего пять условий (Bransford & Jonson, 1972). Второй способ — поместить результаты в таблицу. Для результатов исследования Брэнсфорда и Джонсона можно построить таблицу, подобную табл. 7.2.

Таблица 7.2

Данные исследования Брэнсфорда и Джонсона, представленные в виде таблицы

Таблица 1. Среднее количество запомненных идей как функция различных контекстов для запоминания и воспроизведения материала

Условие	Средняя оценка	Стандартное отклонение
Отсутствие контекста — 1 повторение	3,60	0,64
Отсутствие контекста — 2 повторения	3,80	0.79
Контекст до	8,00	0,65
Контекст после	3,60	0,75
Частичный контекст	4,00	0,60

Примечание. Максимальная возможная оценка равна 14.

Третий способ представления данных — это график. С его помощью можно представить исследование Брэнсфорда и Джонсона, как показано на рис. 7.12. Обратите внимание, что на графике, созданном для экспериментального исследования, зависимая переменная всегда откладывается по вертикальной оси (У), а независимая — по горизонтальной *(X).* Как вы узнаете в следующей главе, ситуация несколько усложняется, если используется более одной независимой переменной. Однако вне зависимости от количества независимых переменных зависимая переменная всегда откладывается по вертикальной оси.

Что использовать, график или таблицу, исследователь решает по своему усмотрению. График может выглядеть очень эффектно, если получены большие различия между результатами или обнаружено взаимодействие (глава 8). Таблицы обычно используют, если данных так много, что график невозможно будет понять, или если исследователь хочет показать читателям точные значения среднего арифметического (в случае графика о точных значениях можно будет лишь догадываться). Единственное правило, которым необходимо руководствоваться, — это то, что одни и те же данные нельзя представлять одновременно и в виде таблицы, и в виде

графика. В целом, данные должны быть представлены таким способом, чтобы результаты, на получение которых вы потратили столько сил, были отображены наиболее ясно и понятно.

Виды графиков

Обратите внимание, что я представил данные исследования Брэнсфорда и Джонсона в виде гистограммы. Почему нельзя сделать это с помощью линейного графика, как на рис. 7.13? В данном случае это не слишком хорошая идея. Проблема заключается в сущности конструкта, использованного в качестве независимой переменной, и непрерывности этой переменной. **Непрерывная переменная** — это переменная, у которой существуют промежуточные значения, а это значит, что она существует в определенном континууме. В качестве примера можно привести дозировку лекарства. В исследовании, в котором сравниваются дозы лекарства в 3,5 и 7 мг, дозировка является непрерывной переменной, ведь мы можем, если потребуется, использовать также 4 или 6 мг. Для изображения результатов в случае непрерывной независимой переменной можно использовать линейный график. Чтобы предположить эффективность промежуточных значений, можно провести интерполяцию по имеющимся точкам и по полученной линии оценить их влияние. В исследовании с лекарствами график может иметь вид, показанный на рис. 7.14. При этом исследователь может быть вполне уверен в оценке эффективности промежуточных значений дозировки, одно из которых помечено звездочкой на рис. 7.15.

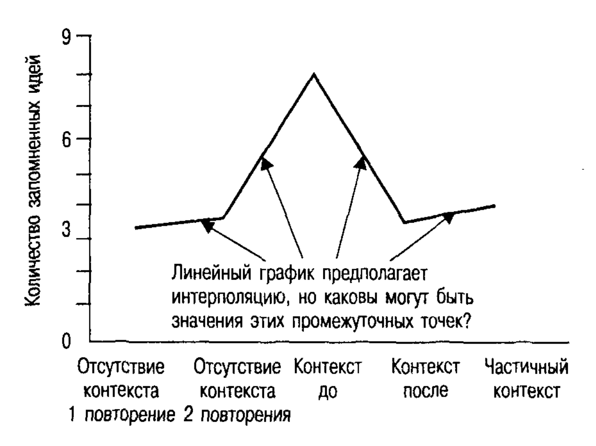


Рис. 7.13. Данные исследования Брэнсфорда и Джонсона, неверно представленные в виде линейного графика

Конечно, если в исследовании используется два значения независимой переменной, довольно сильно отстоящих друг от друга, а зависимость в действительности нелинейная, интерполяция может вызвать проблемы. Так, если в ходе исследования сравниваются дозы лекарства в 2 и 10 мг и получена прямая, изображенная непрерывной линией на рис. 7.16, то если учесть, что истинную зависимость отображает кривая, показанная пунктирной линией, интерполяция эффекта дозы в 5 мг приведет к огромной ошибке. Такое исследование лучше проводить с использованием однофакторного многоуровневого плана.

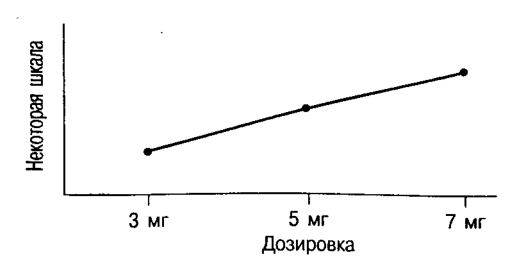


Рис. 7.14. Правильное использование линейного графика в случае непрерывной переменной (дозировка лекарства)

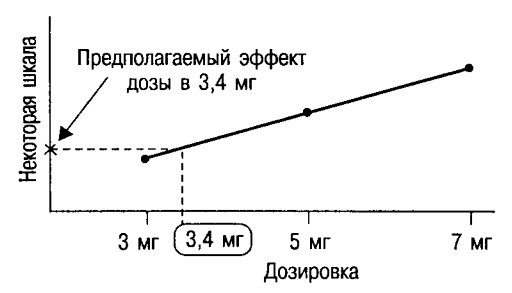


Рис. 7.15. Интерполяция точек линейного графика

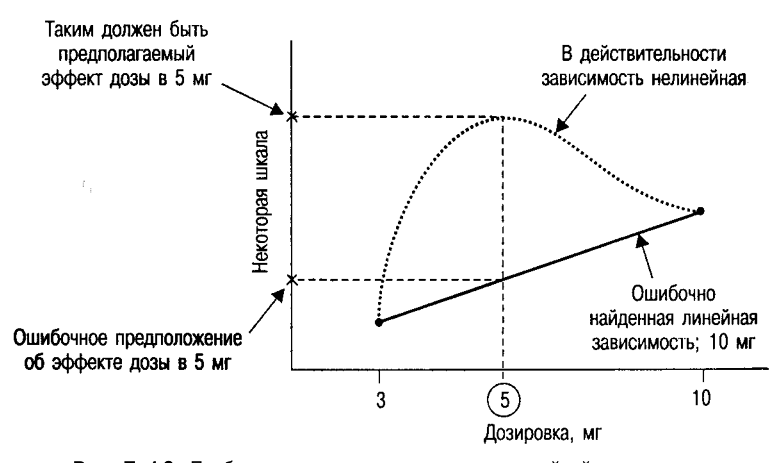


Рис. 7.16. Проблемы интерполяции в случае нелинейной зависимости и большого промежутка

В случае дискретной независимой переменной, каждое значение которой представляет отдельную область, а промежуточных точек просто не существует, ситуация в корне меняется. В таком случае невозможно провести интерполяцию, а следовательно, связать точки с помощью прямой означает предположить существование промежуточных точек, которых в действительности не существует. Поэтому при использовании дискретных переменных, как в исследовании Брэнсфорда и Джонсона (рис. 7.12), обычно строятся гистограммы. Основное правило такое:

Если переменная непрерывна, лучше использовать линейный график, также можно использовать гистограмму.

Если переменная дискретна, лучше использовать гистограмму, линейный график использовать нельзя.

В целом, гистограмму можно использовать как для непрерывных, так и для дискретных данных, а линейный график можно использовать только для непрерывных данных. Обратитесь ко вставке 4.3 — она напомнит вам об этических аспектах представления данных. Легко дезинформировать наивных читателей отчета об исследовании, изменив расстояния на шкале *Y*. Вы как исследователи ответственны за то, чтобы честно представить результаты эксперимента и использовать для этого способ, позволяющий наилучшим образом проиллюстрировать полученные данные.

Анализ однофакторного многоуровневого плана

Как вы уже знаете, если в случае однофакторного двухуровневого плана независимая переменная измеряется с помощью интервальной шкалы или шкалы отношений, для проверки нулевой гипотезы можно использовать коэффициент Стьюдента. Вы можете подумать, что для многоуровневого плана, как, например, в исследовании Брэнсфорда и Джонсона, достаточно будет найти коэффициенты Стьюдента для всех пар условий (например, для контекста до и контекста после). К сожалению, все не так просто. Трудность заключается в том, что проведение многократных проверок по критерию Стьюдента увеличивает риск возникновения ошибки 1-го рода. Чем больше подобных проверок вы проводите, тем больше вероятность ошибочно обнаружить значимые различия. Чтобы охватить все пары условий в исследовании Брэнсфорда и Джонсона, потребуется найти 10 коэффициентов Стьюдента.

Вероятность сделать по крайней мере одну ошибку 1 -го рода при проведении многократных проверок по критерию Стьюдента можно оценить по формуле

$$1 - (1 - a)^{c},$$

где с — это количество проведенных сравнений.

Таким образом, если для исследования Брэнсфорда и Джонсона найти все возможные коэффициенты Стьюдента, то возникнет очень высокая вероятность (4 из 10) сделать но крайней мере одну ошибку 1-го типа:

$$1 - (1 - 0{,}05)^{10} = 1 - (0{,}95)^{10} - 1 - 0{,}60 - 0{,}40.$$

Чтобы избежать проблем, связанных с проведением многократных проверок по критерию Стьюдента в случае однофакторных планов, исследователи используют особую процедуру, носящую название «метод *ANOVA*» (ANalysis Of VAriance — дисперсионный анализ). «Однофакторный» означает наличие одной независимой переменной. По сути, однофакторный метод *ANOVA* проверяет наличие некоторого «общего» значения среди различных значений независимой переменной Так, в исследовании с тремя уровнями нулевая гипотеза будет следующая: «уровень 1 = ^уровень 2 = уровень 3». Однако отклонить нулевую гипотезу еще не значит понять, какой из знаков равенства использован ошибочно. Чтобы точно определить

общее значение, необходимо провести так называемую «последующую проверку» или «анализ *post hoc*» (после факта). В ходе последующей проверки для исследования с тремя уровнями после того, как общий анализ *ANOVA* обнаружит наличие такого значения, будет проведен анализ всех трех сравниваемых пар. Если *ANOVA* не находит общего значения, последующая проверка обычно не производится — ее используют только в том случае, если в дальнейшем возникнут определенные предположения насчет конкретной пары условий. В приложении *С* показан однофакторный метод *ANOVA* и распространенный вариант последующего тестирования, названный *«HSD-тест»* Хаки.

В ходе однофакторного метода *ANOVA* вычисляется «оценка *F*», или «отношение *F*». Хак же, как коэффициент Стьюдента, отношение *F* показывает, насколько вероятно то, что найденные различия в значениях среднего арифметического вызваны случайностью или связаны с влиянием других факторов (возможно, независимой переменной). Метод *ANOVA* очень широко применяется психологами, и если на занятиях по статистике вы еще не успели подробно познакомиться с ним, изучите примеры, приведенные в приложении *С*. Хакже необходимо знать, что обычно, если независимая переменная принимает всего два значения, используется проверка по критерию Стьюдента. Но этот способ не единственный — в такой ситуации также можно применить однофакторный метод *ANOVA*. По сути, проверку по критерию Стьюдента можно рассматривать как особый случай метода *ANOVA*, применяемый, когда независимая переменная принимает два значения.

Все планы, рассмотренные в данной главе, имеют одну общую особенность — наличие одной независимой переменной. В главе 8 будет сделан следующий логический шаг и рассмотрены планы с несколькими независимыми переменными — «факторные планы».

Резюме

Один фактор — два уровня

Простейший экспериментальный план содержит одну независимую переменную, принимающую два значения (два уровня). Переменная в таком плане может быть межсубъектной или внутрисубъектной. Межсубъектными переменными можно управлять непосредственно или их можно отбирать как субъективные факторы. Если переменная управляемая, испытуемых можно распределить по группам случайным образом (план с независимыми группами) или уравнять по переменной, являющейся потенциальным осложнителем, а затем случайным образом распределить (план с уравненными группами). В случае субъективной переменной используется план с неэквивалентными группами. Однофакторные планы с внутрисубъектной переменной иногда называют планами с повторяемыми измерениями (как в случае известного исследования Струпа). Для статистической оценки исследований с двумя уровнями независимой переменной обычно проводят проверку по критерию Стьюдента (при условии данных, полученных с помощью интервальной шкалы или шкалы отношений).

Планы с контрольными группами

Планы с контрольными группами подразумевают наличие хотя бы одного условия, при котором не осуществляется экспериментальное воздействие. Есть следующие разновидности контрольных групп: группы плацебо, часто используемые при исследовании лекарств; группы листа ожидания, которые используют при оценке эффективности определенной программы или терапии; а также эквивалентные контрольные группы, участники которых тщательно уравнены с испытуемыми из экспериментальной группы по определенному фактору, требующему неотрывного контроля.

Один фактор — более двух уровней

При сравнении только двух уровней экспериментальной переменной результаты обязательно будут линейными, так как график, построенный по этим результатам, будет содержать всего две точки. Однако некоторые зависимости нелинейны (например, кривая забывания Эббингхауза) — их можно обнаружить, использовав более двух значений независимой переменной. Увеличение количества уровней можно также использовать для проверки и отклонения (фальсификации) альтернативных гипотез. Как и в случае с двумя уровнями, многоуровневые планы могут быть либо межсубъектными, либо внутрисубъектными. Результаты можно наглядно представить с помощью гистограммы, если независимая переменная является дискретной, или в виде линейного графика, если переменная непрерывна. Обычно статистическая оценка исследований, в которых используется более двух значений независимой переменной, проводится с помощью однофакторного дисперсионного анализа *(ANOVA)* (при условии данных, полученных с помощью интервальной шкалы или шкалы отношений).

Задания для повторения

Выбор ответа

1. Чтобы проверить, влияет ли опыт чтения на выполнение задания Струпа, исследовались дети четырех различных возрастов (5, 7, 9 и 11 лет). (Более взрослые дети предположительно имели больший опыт чтения; использовалось только ключевое условие, названное Струпом НЦСр.) Как вы охарактеризуете экспериментальный план?
   1) однофакторный многоуровневый с независимыми группами;
   2) однофакторный с повторяемыми измерениями;
   3) однофакторный многоуровневый с неэквивалентными группами;
   4) однофакторный двухуровневый с уравненными группами.
2. Что общего имеют все однофакторные планы с повторяемыми измерениями?
   1) они всегда включают контрольную группу;
   2) участники исследуются при каждом из экспериментальных условий более одного раза;

3) наиболее предпочтительным методом создания эквивалентных групп является уравнивание;
4) каждый участник исследуется при каждом из экспериментальных условий.

3. Какой план иллюстрируется примером, посвященным изучению влияния депривации сна на особенности ответов на наводящие вопросы?
   1) внутрисубъектный многоуровневый;
   2) с независимыми группами;
   3) с уравненными группами;
   4) с неэквивалентными группами.
4. Важнейшая особенность сцепленных контрольных групп состоит в том, что,
   1) члены этих групп получают плацебо;
   2) на испытуемых в этих группах оказывается такое же воздействие, как на членов экспериментальных групп;
   3) по сравнению с членами экспериментальных групп испытуемые в этих группах исследуются при меньших значениях независимой переменной;
   4) они используются, если исследователь хочет обнаружить нелинейную зависимость.
5. Предположим, вы пытаетесь повторить исследование памяти, проведенное Эббингхаузом. В понедельник у вас уходит 20 минут на запоминание списка CVC, а во вторник — 15 минут на его повторение. Какова оценка экономии?
   1) 25%;
   2) 75%;
   3) 5%;
   4) ее невозможно вычислить, не зная количество слогов в списке.

Короткие эссе

1. Что общего имеют и чем различаются планы с независимыми группами, планы с уравненными группами и планы с неэквивалентными группами?
2. Почему в примере, посвященном сравнению людей, подвергавшихся и не подвергавшихся депривации сна, был использован план с уравненными группами, а не план с независимыми группами? Какова была переменная уравнивания?
3. Опишите эффект Струпа и использованный автором экспериментальный план.
4. Опишите две разновидности проверки по критерию Стьюдента и на примере планов, изученных в начале данной главы (однофакторные многоуровневые), расскажите о возможностях их использования.
5. С помощью примера, посвященного влиянию алкоголя на скорость реакции, объясните, для чего используются контрольные группы плацебо.
6. На примере исследования «бессознательно воспринимаемых записей» покажите, для чего используются контрольные группы листа ожидания.

7. На примере исследования Вейсса, посвященного изучению взаимосвязей между управлением и стрессом, объясните, для чего нужны эквивалентные контрольные группы.
8. На примере гипотетического исследования влияния кофеина на скорость реакции покажите, как, используя многоуровневые планы, можно обнаружить нелинейный эффект.
9. На примере эксперимента Брэнсфорда и Джонсона, посвященного изучению влияния контекста на запоминание, покажите преимущества использования более двух значений независимой переменной.
10. Опишите, когда лучше всего использовать: а) линейный график, б) гистограмму. Объясните, почему линейный график нельзя использовать в исследовании, в котором сравнивается скорость реакции у женщин и мужчин.

Упражнения

Упражнение 7.1. Определения вида экспериментального плана

Для каждого из описанных ниже исследований укажите независимую и зависимую переменные, особенности первой (межсубъектная или внутрисубъектная; управляемая или субъективная) и назовите, какой экспериментальный план был использован.

1. В исследовании, посвященном изучению влияния булимии на восприятие размера тела, две группы женщин одного возраста (члены одной из групп страдали булимией, а второй — нет) рассматривают серию рисунков, изображающих женщин разного размера, и указывают, какой размер более всего соответствует их восприятию собственного тела.
2. Студентов колледжа, участвующих в исследовании когнитивных карт, попросили с помощью прибора для определения направления точно указать направления, в которых находятся три объекта, не видимые из лаборатории и располагающиеся на разных расстояниях от нее.
3. Три группы дошкольников (случайным образом распределенные по 50 человек в группу) участвуют в исследовании, посвященном изучению настойчивости при выполнении заданий, в котором варьируется время до получения вознаграждения. Детям во всех трех группах раздали трудные головоломки и попросили собирать их, пока не надоест. Одной группе сказали, что по окончании работы все получат по 5 долларов. Вторая группа получит 5 долларов через два дня после окончания эксперимента, а третья — через 4 дня.
4. Для изучения воздействия тесноты на решение задач участников попросили решить серию словесных головоломок, находясь при этом либо в больших, либо в маленьких комнатах. Чтобы получить одинаковое среднее значение вербального *IQ* в группах, исследователи измерили вербальный интеллект участников, а затем распределили их по двум условиям.

Упражнение 7.2. Результаты

Для каждого из приведенных ниже исследований определите, каким образом представить результаты: в виде линейного графика или гистограммы, а затем создайте график, точно отображающий результаты.

1. В исследовании, посвященном изучению влияния марихуаны на непосредственное запоминание списка слов, участники случайным образом распределены на три группы: экспериментальную группу, контрольную группу плацебо и обычную контрольную группу.

   Результат А. Марихуана снижает запоминание, а ожидание действия марихуаны не влияет на запоминание.

   Результат В. Марихуана снижает запоминание, но ожидание действия марихуаны также снижает запоминание.

   Результат С. Кажущееся неблагоприятное действие марихуаны на запоминание связано только с эффектом плацебо.

2. С помощью надежного и валидного теста исследователь оценивает уровень независимости у трех групп студенток, которые провели в колледже 2 месяца. Высокий уровень независимости означает возможность справляться с делами самостоятельно. Одна группа (О300) состоит из студенток, живущих в общежитии, чей дом находится за 300 или более миль от студенческого городка; студентки из второй группы (О100) живут в общежитии, а их родители живут на расстоянии менее 100 миль от студенческого городка; студентки из третьей группы живут дома (Д).

   Результат А. Живущие дома студентки более независимы, чем проживающие в общежитии.

   Результат В. Чем дальше дом от студенческого городка, тем более независимым является человек.

   Результат С. Живущие дома студентки и члены группы О300 очень независимы, а студентки из группы О100 — нет.

3. Животные изучают лабиринт, и в процессе этого фиксируются совершаемые ими ошибки (т. е. неправильные повороты). В конце каждой попытки, дойдя до цели, они получают пищевое вознаграждение. Одной группе крыс еду выдают сразу по достижении цели (задержка 0). Второй группе еду выдают через 5 секунд после достижения цели (задержка 5 секунд).

   Результат А. Задержка подкрепления затрудняет обучение.

   Результат В. Задержка подкрепления не влияет на обучение.

4. Игроки в баскетбол совершают три серии по 20 свободных бросков в кольцо, испытывая при этом возбуждение различной силы: слабое, среднее или сильное. Предполагается, что слабое возбуждение возникает, если неудачный бросок наказывается пробежкой вокруг игрового поля (т. е. наказание минимально и не вызывает особого возбуждения). Среднее возбуждение означает два круга вокруг поля, а сильное — четыре круга (т. е. достаточно большое наказание, чтобы создать сильное возбуждение, скорее всего, в виде тревожности). Используется план с повторяемыми измерениями; проведено позиционное уравнивание.

   Результат А. Существует линейная зависимость между возбуждением и выполнением задания: рост возбуждения снижает качество выполнения задания.

   Результат В. Существует нелинейная зависимость между возбуждением и выполнением задания: задание выполняется хорошо только при среднем уровне возбуждения.

ГЛАВА 8

Экспериментальный план II факторные планы

Обзор задач главы

В главе 7 вы познакомились с основными видами экспериментальных планов с одной независимой переменной, которая принимает одно или более значений. Следующий логический шаг — увеличение числа независимых переменных и переход к факторным планам. Этому и посвящена данная глава. Изучив ее, вы:

- сможете описывать факторные планы с помощью стандартизованной системы обозначений (2 х 2, 3 х 5 и т. д.);
- научитесь правильно размещать данные в факторной матрице;
- поймете смысл явления основного эффекта и научитесь обнаруживать его.
- поймете смысл явления эффекта взаимодействия и научитесь обнаруживать его;
- научитесь определять разновидности факторных планов, соответствующие однофакторным планам из главы 7 (с независимыми, уравненными, неэквивалентными группами или с повторяемыми измерениями);
- сумеете опознать смешанные планы и поймете, почему они не всегда требуют применения позиционного уравнивания;
- научитесь определять факторный план Р х Е, узнаете о его связи с фигурой Курта Левина и поймете, что означает основной эффект и эффект взаимодействия для такого плана;
- научитесь вычислять количество участников, необходимое для каждой разновидности факторного плана;
- поймете, как стать этически компетентным экспериментатором.

Во время изучения курса исследовательских методов вы, возможно, заметили, что психологи-экспериментаторы иногда пользуются особым, понятным только им, языком. Они рассуждают о операционализации конструктов, применении метода стебля и листа и устранении осложнителей. Говоря о регрессии, они не имеют в виду теорию Фрейда. Но это только начало — изучив данную главу, вы сможете сказать, например, следующее: «Это был смешанный факторный план два на три, с одним основным эффектом у многократно измеряемой переменной и взаимодействием». Начнем с основ.

Основы факторных планов

Предположим, вас интересует проблема памяти и вы хотите выяснить, можно ли улучшить эту способность, научившись при запоминании слов формировать зрительные образы. Вы разрабатываете простой эксперимент с двумя группами. Одна группа при запоминании создает зрительные образы, а вторая использует механическое повторение. Допустим, что вы хотите также узнать, как влияет на память скорость, с какой показывается список слов. Вы еще раз проводите простой эксперимент с двумя группами. Одним участникам вы показываете слова со скоростью 2 с/слово, а другим — 4 с/слово. Если использовать факторный план, то *оба* эти эксперимента можно провести в рамках одного исследования.

По определению **факторный план** подразумевает использование более чем одной независимой переменной (которые также называются «факторами»). В принципе, переменных может быть сколько угодно, но на практике обычно используется два или три фактора, реже четыре. Далее, на примере исследования памяти мы познакомимся с системой обозначений, используемой для описания факторных планов.

Обозначение факторных планов

Во-первых, факторный план описывается с помощью системы нумерации, показывающей количество независимых переменных и количество значений (уровней), принимаемых каждой переменной. Так, факторный план 2x3 (читается «два на три») имеет две независимые переменные; первая переменная принимает два значения, а вторая — три. Факторный план 3x4x5 имеет три независимые переменные, принимающие три, четыре и пять значений соответственно. Исследование памяти проводится в соответствии с факторным планом 2x2, где независимые переменные «способ запоминания» (образы или механическое повторение) и «скорость показа» (слово в 2 и 4 с) имеют по два уровня.

Во-вторых, изучаемые в ходе факторного исследования условия можно получить, определив все возможные комбинации уровней каждой независимой переменной. Для исследования памяти можно составить следующую таблицу условий, или **факторную матрицу:**

		Скорость показа	
		2 с/слово	4 с/слово
Способ запоминания	Образы	Образы/2 с	Образы/4 с
	Повторение	Повторение/2 с	Повторение/4 с

Прежде чем продолжить изложение, я хотел бы обратить ваше внимание на один важный момент. До сих пор я использовал термины «условия эксперимента» и «уровни независимой переменной» так, как если бы они означали одно и то же. В случае однофакторных исследований эти понятия действительно взаимозаменяемы, но если используется факторный план, то это не так. Для любых экспериментальных планов термин «уровень» относится к количеству значений, принимаемых независимой переменной. В случае факторного плана термин «условия» соответствует

количеству ячеек матрицы, подобной рассмотренной выше. Таким образом, исследование памяти с планом 2 x 2 содержит *две* независимые переменные, каждая из которых имеет по *два* уровня. План состоит из *четырех* различных условий — по одному на каждую ячейку матрицы. Количество условий для любого факторного плана можно определить, найдя произведение чисел, обозначающих размерность плана. План размером 3 x 3 содержит 9 условий; план, обозначаемый как 2 χ 2 χ 4, — 16.

Обобщенную факторную матрицу размером 2x2 можно наглядно представить следующим образом:

		Фактор B	
		Уровень B1	Уровень B2
Фактор A	Уровень A1	УсловиеA1B1	Условие A1B2
	Уровень A2	УсловиеA2B1	Условие A2B2

Необходимо ясно понимать применяемую систему обозначений, так как при использовании пакетов статистических программ вы встретитесь именно с такими обозначениями. Если компьютер просит вас ввести данные для ячейки A2B1, а вы по ошибке вводите данные для ячейки A1B2, анализ будет проведен и вы получите прекрасную распечатку неверных результатов. Очень важно вводить данные в нужные ячейки. В приложении *C* при изучении метода *ANOVA* вы также встретитесь с этой системой обозначений.

В табл. 8.1 показано, как распределяются факторы и нумеруются ячейки в случае планов 2x4 и 2x2x2. Пока не обращайте внимания на матрицы с закрашенными ячейками — они станут вам понятны после того, как вы изучите следующий раздел.

Таблица 8.1

Модели факторных планов

1. Факторный план 2 × 4:

	B1	B2	B3	B4
A1	A1B1	A1B2	A1B3	A1B4
A2	A2B1	A2B2	A2B3	A2B4

Проверка основного эффекта для A (т. е. сравнение A1 и A2).

	B1	B2	B3	B4
A1	A1B1	A1B2	A1B3	A1B4
A2	A2B1	A2B2	A2B3	A2B4

Проверка основного эффекта для B (т. е. сравнение B1, B2, B3 и B4).

	B1	B2	B3	B4
A1	A1B1	A1B2	A1B3	A1B4
A2	A2B1	A2B2	A2B3	A2B4

2. Факторный план 2 × 2 × 2:

C1

	B1	B2
A1	A1B1C1	A1B2C1
A2	A2B1C1	A2B2C1

C2

	B1	B2
A1	A1B1C2	A1B2C2
A2	A2B1C2	A2B2C2

Проверка основного эффекта для А (т. е. сравнение A1 и A2).

C1

	B1	B2
A1	A1B1C1	A1B2C1
A2	A2B1C1	A2B2C1

C2

	B1	B2
A1	A1B1C2	A1B2C2
A2	A2B1C2	A2B2C2

Проверка основного эффекта для В (т. е. сравнение B1 и B2).

C1

	B1	B2
A1	A1B1C1	A1B2C1
A2	A2B1C1	A2B2C1

C2

	B1	B2
A1	A1B1C2	A1B2C2
A2	A2B1C2	A2B2C2

Проверка основного эффекта для С (т. е. сравнение C1 и C2).

C1

	B1	B2
A1	A1B1C1	A1B2C1
A2	A2B1C1	A2B2C1

C2

	B1	B2
A1	A1B1C2	A1B2C2
A2	A2B1C2	A2B2C2

Результаты: основной эффект и взаимодействие

Факторные исследования дают два вида результатов: основной эффект и взаимодействие. Основной эффект показывает общее влияние независимых переменных, а взаимодействие отражает совместное действие переменных, приводящее к получению более сложных результатов. Рассмотрим оба эти вида результатов более подробно.

Основной эффект

В эксперименте, который мы будем использовать в качестве модели, исследователь изучает действие двух независимых переменных: способа запоминания и скорости показа. В факторных планах для обозначения общего влияния независимой переменной используется термин **основной эффект.** В исследовании с двумя независимыми переменными, например в факторном плане 2x2, может быть не более двух основных эффектов. Для определения основного эффекта одного фактора необходимо использовать все значения другого (других) фактора (факторов). Продемонстрируем это на примере исследования памяти. Чтобы определить основной

эффект способа запоминания, необходимо обобщить данные двух скоростей показа. Другими словами, информацию в ячейках, имеющих более светлый оттенок (образы), следует обобщить и затем сравнить ее с обобщенными данными из более темных ячеек (повторение):

		Скорость показа (B)	
		2 с/слово B1	4 с/слово B2
Способ *запоминания* (*A*)	Образы A1	Образы/2 с A1B1	Образы/4 с A1B2
	Повторение A2	Повторение/2 с A2B1	Повторение/4 с A2B2

Аналогично, чтобы определить основной эффект скорости показа, необходимо обобщить данные двух способов запоминания. В приведенной ниже матрице влияние скорости показа оценивается с помощью сравнения всей информации в более светлых ячейках (2 с/слово) со всеми данными в более темных ячейках (4 с/слово):

		Скорость показа (B)	
		2 с/слово B1	4 с/слово B2
Способ *запоминания* (*A*)	Образы A1	Образы/2 с A1B1	Образы/4 с A1B2
	Повторение A2	Повторение/2 с A2B1	Повторение/4 с A2B2

Теперь еще раз посмотрите на закрашенные матрицы в табл. 8.1— как вы видите, я выделил одним цветом ячейки, которые необходимо обобщать при анализе основного эффекта различных переменных в планах 2 χ 4 и 2 x2 χ 2.

Рассмотрим гипотетические данные, которые могут быть получены в ходе такого эксперимента с памятью. Предположим, при каждом условии (соответствует каждой ячейке) в эксперименте принимают участие 25 испытуемых, а задание заключается в запоминании списка из 30 слов. Среднее количество слов, запомненных участниками при каждом из четырех условий, будет следующим:

		Скорость показа	
		2 с/слово	4 с/слово
Способ *запоминания*	Образы	17	23
	Повторение	12	18

Приводит ли использование визуальных образов к улучшению запоминания в сравнении с механическим повторением? Другими словами, соответствует ли переменной «способ запоминания» основной эффект? Чтобы узнать это, необходи-

мо сравнить все данные из ячеек «образы» со всеми данными из ячеек «повторение». Для этого требуется вычислить значения среднего арифметического по рядам. Среднее арифметическое в ряду «образы» составляет 20 слов [(17 + 23)/2 = = 40/2 = 20], а в ряду «повторение» — 15 слов [(12 + 18/2 = 30/2 = 15)]. Спрашивая о соответствии основного эффекта способу запоминания следует поставить следующий вопрос: «Значима или случайна разница между значениями среднего арифметического, равными 20 и 15?»

Точно таким же образом вычисление значений среднего арифметического по столбцам позволяет узнать, соответствует ли основной эффект скорости показа. Для столбца «2 с/слово» среднее арифметическое составляет 14,5 слов, а для столбца «4 с/слово» — 20,5 слов (обязательно проверьте). Поместим все полученные результаты в одну таблицу:

	Скорость показа		
	2 с/слово	4 с/слово	Итого
Образы	17	23	**20,0**
Повторение	12	18	**15,0**
Итого	**14,5**	**20,5**	

Судя по этим данным, запоминание улучшается при использовании образов (20 > 15), а также при более низкой скорости показа слов (20,5 > 14,5). Таким образом, в данном случае мы обнаружили два основных эффекта (но судить о значимости различий можно, только проведя дисперсионный анализ). В качестве примера эксперимента, имеющего очень сильный основной эффект, рассмотрим реальное исследование влияния визуальных образов на память.

Пример 14. Значимый основной эффект

Использованный ранее гипотетический пример, в котором для улучшения памяти применялись зрительные образы, возник не на пустом месте. Исследователи, занимающиеся когнитивной психологией, уже давно знают, что объем запоминаемой информации можно увеличить, если использовать зрительные образы. Уоллен с соавторами (Wollen, Weber, & Lawry, 1972) провели исследование, в котором попытались по отдельности оценить два фактора, способных увеличить эффективность использования зрительных образов. В одном из более ранних исследований было обнаружено, что для улучшения запоминания следует использовать странные и причудливые, а не обычные образы. В основе этой идеи лежит правило о том, что необычные вещи выделяются в сознании человека. Другое исследование показало, что информация запоминается лучше, если ее отдельные элементы можно объединить в один образ. Уоллен и др. изучали два фактора: причудливость и комбинированность[1]. Участникам выдавали пары слов для запоминания, например

[1] Также изучался третий фактор, названный «связь». Одни пары слов сопровождались иллюстрациями, как показано на рис. 8.1, а изображения, сопровождающие другие пары, не были связаны со словами. Запоминание было лучше при использовании связанных со словами изображений.

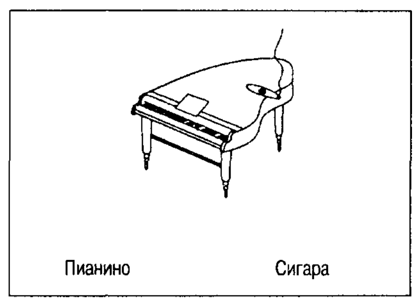
«пианино—сигара». Если в ходе опроса испытуемым показывали слово «пианино», они должны были ответить «сигара».

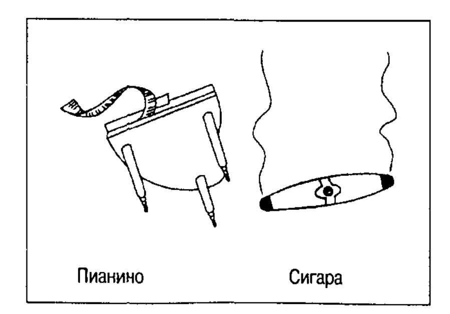
В процессе демонстрации пар слов участникам показывали один из четырех рисунков, включая изображенные на рис. 8.1 для пары «пианино—сигара». На первых двух рисунках предметы скомбинированы, а на вторых двух — нет; кроме того, два рисунка являются необычными (т. е. причудливыми), а два —обычными. Среднее число пар слов, запомненных при каждом из четырех условий факторного плана 2 x 2 с независимыми группами, показало сильный основной эффект для фактора «комбинированность», но отсутствие значимого влияния фактора «причудливость» (максимальная оценка = 9):

	Причудливые	Обычные	Итого
Комбинированные	6,67	6,60	6,64
Не комбинированные	3,05	3,50	3,28
Итого	4,86	5,05	

Таким образом, фактор «причудливости» не играет никакой роли (по крайней мере в данном примере) — общее количество запомненных слов было практически одинаковым при использовании причудливых (4,86 из 9) и обычных (5,05) изображений. Комбинирование изображений, в свою очередь, имело значительный эффект. Количество запомненных слов при комбинировании была выше (6,64), чем при его отсутствии (3,28). На рис. 8.2 полученные результаты представлены в виде гистограммы.







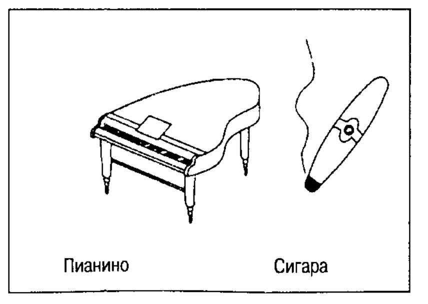


Рис. 8.1. Пример стимульного материала из исследования зрительных образов, проведенного Уолленом и др. (1972)

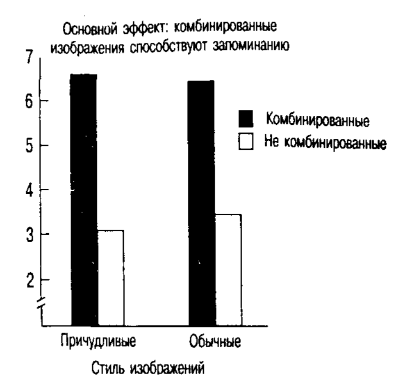


Рис. 8.2. Гистограмма, отображающая основной эффект для комбинированных **изображений, отсутствие** основного эффекта для их «причудливости» и отсутствие взаимодействия (построена на основании данных Уоллена и др., 1972)

Основной эффект — это важная особенность факторных планов, но заметное преимущество таких планов перед однофакторными заключается в возможности показать эффект взаимодействия. В эксперименте, построенном по факторному плану, демонстрируется взаимодействие, если влияние одной независимой переменной зависит от значения другой независимой переменной. Идею взаимодействия довольно сложно понять, но важно это сделать, поскольку наличие взаимодействий нередко приводит к получению самых интересных результатов для факторных исследований. Для начала рассмотрим простой пример. Предположим, я считаю, что общую психологию лучше преподавать как лабораторный самостоятельный курс, а не как обычный лекционный, но хочу узнать, верно ли это для всех студентов или только для некоторых. Возможно, что для студентов, специализирующихся в естественных науках, лабораторные занятия будут более полезны. Для проверки этой идеи мне необходимо сравнить лабораторный и лекционный курсы по общей психологии, а также различные типы студентов, например специализирующихся в гуманитарных и естественных науках. Для этого потребуется план 2x2, который можно представить следующим образом:

		Вид занятий	
		В основном лабораторные	В основном лекционные
Специализация студентов	Естественные науки	Студенты-естественники на лабораторных занятиях	Студенты-естественники на лекционных занятиях
	Гуманитарные науки	Студенты-гуманитарии на лабораторных занятиях	Студенты-гуманитарии на лекционных занятиях

В подобном исследовании зависимая переменная должна показывать уровень знаний — мы будем рассматривать оценки от 1 до 100, полученные по стандартизованному тесту знаний по общей психологии, который используют во время сессии. Допустим, что получены следующие результаты:

	В основном лабораторные	В основном лекционные
Естественные науки	80	70
Гуманитарные науки	70	80

Говорят ли такие данные о наличии основного эффекта? Нет, так как все значения среднего арифметического по рядам и столбцам одинаковы: 75. Означает ли это, что в ходе исследования вообще ничего не произошло? Нет, не означает. Студенты-естественники показали лучшие результаты по лабораторным занятиям, а студенты-гуманитарии — по лекционным. Выражаясь в терминах определения взаимодействия, влияние одной переменной (вид занятий) зависит от значения другой переменной (специализация). Таким образом, даже если основной эффект отсутствует, может существовать взаимодействие, приводящее к интересным результатам.

Пример с изучением общей психологии отражает также заметное преимущество факторных планов перед однофакторными. Предположим, что вы провели однофакторное двухуровневое исследование по сравнению лабораторного и лекционного курсов по общей психологии. Скорее всего, вы используете для этого план с уравненными группами, общие баллы студентов и специализацию как переменную уравнивания. Вы можете привлечь тех же студентов, что и в примере с факторным планом, но в случае однофакторного исследования будут получены следующие результаты:

Лабораторные занятия: 75 Лекционные занятия: 75

Получив подобные данные, вы можете заключить, что использование лабораторных занятий по общей психологии не дает никаких результатов. Однако факторный план покажет вам, что результаты есть, но только для определенного типа студентов. Говоря кратко, факторный план может быть более информативным, чем однофакторный. Чтобы более подробно проиллюстрировать понятие взаимодействия, рассмотрим еще один пример.

Пример 15. Взаимодействие при отсутствии основного эффекта

Существует заслуживающее внимания исследование, показавшее, что людям легче вспомнить некоторое событие, если они находятся в том же месте или той же ситуации, в которой это событие встретилось им в первый раз. Возможно, вы и сами сталкивались с такой ситуацией, когда могли найти потерянные ключи, поместив себя мысленно или физически в то место, в котором вы видели их в последний раз.

Это исследование, получившее награду за самую креативную проверку гипотезы связи контекста с запоминанием, было проведено Годденом и Бэддели (Godden & Baddely, 1975). Авторы использовали факторный план 2 χ 2, в котором испытуемые изучали список из 36 слов в одних условиях, а затем воспроизводили его в тех же или других условиях. Креативность эксперимента определялась выбором условий. Испытуемыми были члены клуба ныряльщиков, которые запоминали слова либо на берегу, либо в воде на глубине в 20 футов! Первой независимой переменной было место, в котором проходило изучение слов. Она принимала два значения:

«на земле» и «под водой». Второй переменной было место, в котором происходило воспроизведение слов. Она также имела два уровня — наземный и подводный. Таким образом, исследование включало четыре условия:

1) изучение на земле — воспроизведение на земле;
2) изучение на земле — воспроизведение под водой;
3) изучение под ведой — воспроизведение на земле;
4) изучение под водой — воспроизведение под водой.

Все ныряльщики последовательно проходили все четыре условия, следовательно, план был факторным с повторяемыми измерениями. Были получены следующие результаты, отражающие среднее количество запомненных слов:

		Где проходило воспроизведение		
		На земле	Под водой	Итого
Где проходило запоминание	На земле	13,5	8,6	11,1
	Под водой	8,4	14,4	9,9
	Итого	11,0	10,0	

Эти данные похожи на результаты гипотетического исследования вариантов обучения общей психологии студентов-естественников и гуманитариев. Значения среднего арифметического по рядам и столбцам практически одинаковы, что говорит об отсутствии основного эффекта. Следовательно, для эксперимента в целом не имеет значения, где проходят запоминание (нет значимых различий между 11,1 и 9,9) и воспроизведение (11,0 практически не отличается от 10,0) слов. Но внимательное рассмотрение всех четырех значений среднего арифметического показывает наличие взаимодействия. Когда ныряльщики запоминали слова на земле, они успешно вспоминали их на земле (13,5), но не в воде (8,6); изучая слова под водой, они с трудом вспоминали их на земле (8,4), но неплохо справлялись с заданием под водой (11,4). Это означает, что изучение было наиболее успешным, когда контекст при запоминании и воспроизведении был один и тот же. На рис. 8.3 данные представлены в виде гистограммы.

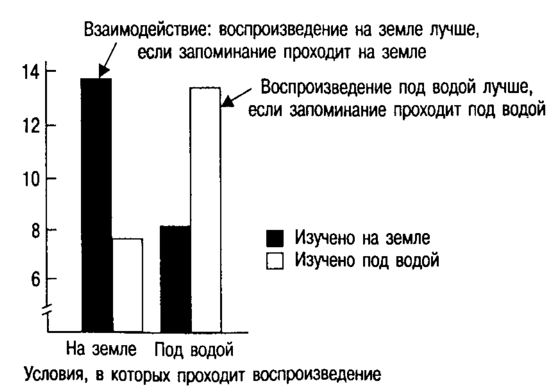


Рис. 8.3. Гистограмма, отображающая взаимодействие между условиями, в которых проходят запоминание и воспроизведение

Вас могут заинтересовать процедуры контроля, разработанные для этого исследования. Из-за шума дыхания ныряльщики плохо слышали слова, записанные на магнитофон и передаваемые через устройство «для подводной связи с ныряльщиками», а поэтому они были вынуждены задерживать дыхание на время до 6 секунд, пока звучали группы по три слова в каждой. Между этими группами были «дыхательные» паузы в четыре секунды. Чтобы сделать дыхательный режим запоминания одинаковым для всех экспериментальных условий, данная процедура имитировалась при изучении слов на земле. Кроме того, проверка запоминания проводилась после того, как ныряльщики завершали свою ежедневную работу. Таким образом, «испытуемые начинали каждую сессию, находясь приблизительно в одинаковом состоянии: мокрые и замерзшие» (Godden & Baddeley, 1975, p. 327). Если вас заинтересовало слово «замерзшие», так как вы думаете, что у британских исследователей был хороший повод провести свои эксперименты на Багамах, то вы ошибаетесь, они не были *настолько* креативными (или не имели достаточного финансирования). Исследование было проведено на западном берегу *Шотландии*.

Сочетание основного эффекта и взаимодействия

Пример, показывающий зависимость памяти от контекста, иллюстрирует лишь один вид результатов, возможных в случае факторного плана (наличие взаимодействия, но отсутствие основного эффекта). Существуют многие другие варианты. Например, в случае простого плана размером 2x2 есть восемь возможностей:

1) основной эффект только для фактора Л;
2) основной эффект только для фактора *В;*
3) основные эффекты только для факторов *А* и *В;*
4) основной эффект для фактора *А* плюс взаимодействие;
5) основной эффект для фактора *В* плюс взаимодействие;
6) основные эффекты для факторов *АиВ* плюс взаимодействие;
7) только взаимодействие, отсутствие основных эффектов;
8) отсутствие и основных эффектов, и взаимодействия.

Рассмотрим коротко некоторые их этих результатов на примере гипотетического эксперимента с двумя переменными: способом запоминания и скоростью показа. Для каждого из следующих примеров я подобрал данные, которое могли быть получены в ходе исследования влияния способа запоминания и скорости показа на запоминание списка из 30 слов, представил эти данные в виде графика и привел устное объяснение результатов. Я не стал рассматривать все восемь описанных выше возможностей, а привел лишь те результаты, которые действительно могли быть получены в подобном исследовании.

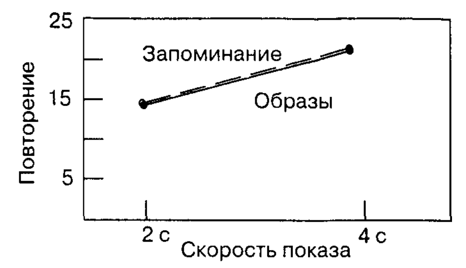
1. Создание зрительных образов улучшает запоминание вне зависимости от скорости показа, которая не влияет на запоминание. Это означает наличие основного эффекта для фактора *А* (способ запоминания) и его отсутствие для скорости показа *(В).* Данный график должен напомнить вам график из примера 12 (рис. 8.2) с одним отличием: он линейный.

	2 с	4 с	Итого
Образы	22	22	22
Повторение	14	14	14
Итого	18	18	

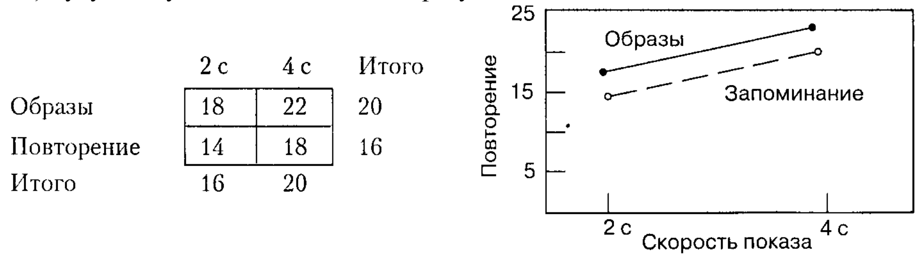
25
15
5
Повторение
Образы
Запоминание
2 с
Скорость показа
4 с

2. Запоминание улучшается при более медленном показе слов, а способ запоминания не влияет на него. Это означает наличие основного эффекта только для фактора *B* (скорость показа).

	2 с	4 с	Итого
Образы	14	22	18
Повторение	14	22	18
Итого	14	22	



3. Запоминание улучшается при более низкой скорости показа, кроме того, на него влияет способ запоминания. В данном случае основной эффект имеется у обоих факторов. При проведении подобного исследования, скорее всего, будут получены именно такие результаты.

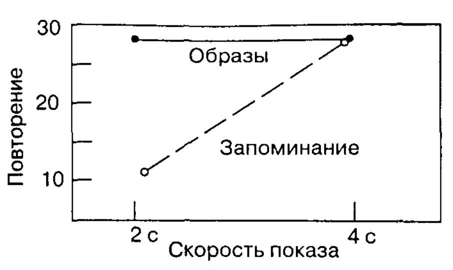


	2 с	4 с	Итого
Образы	18	22	20
Повторение	14	18	16
Итого	16	20	

4. При показе каждого слова по 2 с использование зрительных образов заметно улучшает запоминание (с 12 до 28), а при скорости показа 4 с/слово запоминание почти абсолютное вне зависимости от способа. Это означает наличие взаимодействия между способом запоминания и скоростью показа. В данном случае на взаимодействие мог оказать влияние так называемый **эффект потолка,** возникающий, если оценки для различных условий настолько близки к максимуму (30 слов в этом примере), что различий между ними не возникает[1]. В нашем примере группа, создававшая визуальные образы, запомнила почти все слова вне зависимости от скорости показа. Чтобы проверить наличие эффекта потолка, вы можете повторить исследование со списком из 50 слов и посмотреть, улучшается ли запоминание для группы образы/4 с.

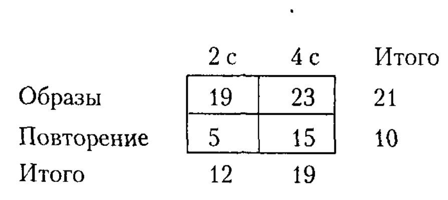
Также оценки при двух условиях могут быть одинаковыми из-за того, что они достигли минимума. Да, действительно, это называется *эффект пола.*

	2 с	4 с	Итого
Образы	28	28	28
Повторение	12	28	20
Итого	20	28	



Вас, вероятно, удивило явное присутствие основного эффекта в этом примере. И действительно, значения среднего арифметического по рядам (20 и 28) и столбцам (также 20 и 28) показывают наличие значимого общего влияния обоих факторов. Формально анализ, скорее всего, покажет присутствие статистически значимых основных эффектов в данном примере, но это говорит лишь о том, что при интерпретации результатов взаимодействие обычно пользуется приоритетом над основным эффектом. В данном случае основные эффекты не имеют особого значения, а утверждение, что визуальные образы приводят к общему улучшению запоминания, не вполне правильно. В действительности они способствуют запоминанию только при более высокой скорости показа. Также будет неверным вывод о том, что при скорости показа 4 с на слово запоминание лучше, чем при скорости 2 с на слово, — это так только для группы, использующей механическое повторение. Таким образом, главным здесь является взаимодействие. Если в факторном исследовании обнаружены и взаимодействия, и основные эффекты, то взаимодействия должны интерпретироваться в первую очередь.

5. Но это не значит, что основной эффект не играет никакой роли при наличии взаимодействия. Рассмотрим следующий пример:



	2 с	4 с	Итого
Образы	19	23	21
Повторение	5	15	10
Итого	12	19	



В этом случае использование зрительных образов в целом способствует запоминанию (т. е. есть основной эффект для *А:* 21 > 10). Но более низкая скорость показа также улучшает запоминание в обеих группах (т. е. также есть основной эффект для *В:* 19 > 12). Необходимо отметить оба эти результата. Взаимодействие показывает, что снижение скорости показа несколько улучшает запоминание для группы, использующей зрительные образы (23 немного лучше, чем 19), но имеет значительный эффект для контрольной группы (15 намного лучше, чем 5). Другими словами, при более высокой скорости показа использование зрительных образов особенно эффективно (19 заметно лучше, чем 5). При более низкой скорости показа зрительные образы улучшают запоминание, но не так сильно, как при более высокой (23 немного лучше, чем 15).

Изучая графики, вы, возможно, обратили внимание на одну характерную особенность взаимодействий. В целом, если прямые графика параллельны друг другу, то взаимодействия не существует. Если прямые не параллельны, велика вероятность взаимодействия. Конечно, это правило о параллельности прямых является лишь общим руководством и наличие взаимодействия (а по сути, непараллельность прямых) должен выявлять дисперсионный анализ.

Обнаружить взаимодействие по параллельности прямых легче при использовании линейного графика, чем в случае гистограммы. Таким образом, отмеченное в главе 7 общее правило о том, что линейные графики используются только при непрерывных переменных, в ситуации, когда наиболее важным является наличие взаимодействия, нередко игнорируется. Например, в исследовании Келтнера с соавторами (Keltner, Ellsworth, & Edwards, 1993), когда участников попросили оценить вероятность возникновения несчастного случая (автомобильной аварии), было обнаружено взаимодействие между эмоциями, испытываемыми участниками во время эксперимента, и имеющейся у них информацией о том, чем вызвано событие, людьми или стечением обстоятельств. Если участники были грустными, они считали, что вероятность происшествия по вине обстоятельств (например, мокрая дорога) выше, чем по вине человека (например, недостатки навыков вождения). Если участники испытывали злость, результат был противоположным. Как показано на рис. 8.4, несмотря на то что по оси *X* откладывается дискретная переменная, был использован линейный график. Вероятно, Келтнер и его соавторы (1993) хотели как можно яснее показать взаимодействие и поэтому проигнорировали правило о дискретных переменных. Я хочу еще раз повторить сделанное ранее замечание и сказать, что для представления данных самое важное — сделать полученные напряженным трудом результаты как можно более понятными для читателей.

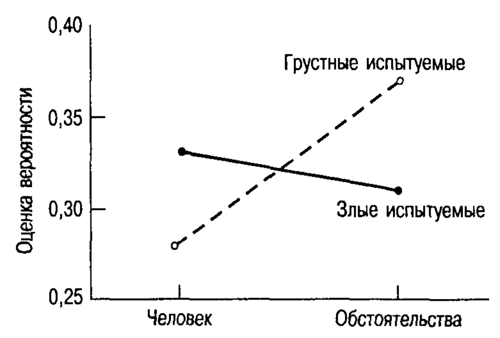


Рис. 8.4. Использование линейного графика для отображения взаимодействия (Kelneretal., 1993)

Прежде чем приступать к изучению классификации факторных планов, прочитайте вставку 8.1. В ней описывается один из самых известных психологических экспериментов, подтвердивший идею о том, что между последними усилиями по подготовке к экзамену и моментом сдачи экзамена человеку необходим сон. Эксперимент был проведен в начале 20-х гг. прошлого века, когда термин «факторный план» в области экспериментальной психологии еще не вошел в употребление,

а дисперсионный анализ — самый распространенный в статистике способ анализа факторных планов — находился в стадии формирования. Это исследование иллюстрирует тенденцию, которая привела к разработке факторных планов, — желание ученых одновременно изучать более одной независимой переменной.

ВСТАВКА 8.1

Классические исследования — запоминание во сне

Как вы уже знаете, основное достоинство факторных планов состоит в том, что с их помощью можно одновременно изучать влияние двух или более независимых переменных, а также взаимодействие этих переменных. Хотя ни термин «факторный план», ни статистические методы анализа факторных планов до Второй мировой войны практически не использовались (см. вставку 8.3), попытки одновременно изучать более одной переменной совершались задолго до этого времени. Классический пример этого представляет собой исследование Джона Дженкинса и Карла Далленбаха, опубликованное вместе с другими исследованиями, проведенными Корнельским университетом в 1924 г. Это исследование до сих пор можно встретить во многих учебниках по общей психологии в качестве типичного примера «ретроактивного влияния»- особенности памяти, заключающейся в том, что если между моментами запоминания и воспроизведения информации вклинивается какая-либо другая умственная деятельность, то возможности памяти снижаются. Использовался факторный план 2 χ 4 с повторяемыми измерениями. «2» говорит о том, что умственная деятельность либо разделяла, либо нет изучение и воспроизведение, а «4» означает использование четырех различных интервалов хранения информации в памяти: правильность воспроизведения проверяли через 1,2,4 или 8 часов после запоминания. Интересным (а соответственно и известным) это исследование является потому, что в нем рассматривается фактор, имеющий два уровня. Между запоминанием и воспроизведением участники исследования либо бодрствовали и занимались обычными студенческими делами, либо спали в Корнельской психологической лаборатории. Эксперименты подтвердили предположение о том, что сон снижает степень ретроактивного влияния, а следовательно, улучшает воспроизведение информации.

Внимательное рассмотрение этого исследования обнаруживает черты, типичные для экспериментов 20-х гг., а также показывает, что исследователи подошли к вопросам контроля за экспериментом так же внимательно, как это делают в наше время. Во-первых, как вы узнаете в главе 10, на заре психологии исследования проводились с привлечением очень небольшого числа участников. В противоположность тому, как организуются исследования памяти сегодня, когда набирается много участников, а затем проводится статистическая обработка данных, первые эксперименты проводились с участием лишь одного, двух или трех испытуемых с последующим отчетом о результатах по всем участникам. Для повторного проведения экспериментов участники набирались дополнительно. Именно так было организовано исследование Дженкинса и Далленбаха (Jenkins & Dallenbach, 1924): в нем участвовало лишь два испытуемых (называемых «наблюдатели» - еще одна особенность исследований того времени), оба - старшекурсники Корнельского университета. При небольшом количестве участников исследователи старались получить от них как можно больше информации, поэтому в данном исследовании был использован, выражаясь современным языком, план с повторяемыми измерениями. Это означает, что для всех восьми ячеек плана 2 x 4 данные были получены от обоих студентов. Студенты должны были учить, а затем воспроизводить списки слогов по восемь раз для каждого из восьми экспериментальных условий - всего 64 попытки (из-за проблем с расписанием пришлось отменить две попытки с бодрствованием для одно-

го студента и одну - для другого). Если вы считаете, что такое исследование было слишком тяжелым для двух старшекурсников из Корнелла, то вы абсолютно правы. Двое студентов и Далленбах, выполняющий функции экспериментатора, «в ходе проведения экспериментов жили в лаборатории» в комнате, временно приспособленной под спальню (р. 606). За исключением тех дней, когда студенты были «не расположены» к работе, эксперименты проводились каждый день и каждую ночь с 14 апреля по 7 июня 1923 г. Я не знаю, когда в том году у студентов Корнелла был выпускной вечер, но я думаю, что сегодня немногие студенты согласятся отдать науке последние полтора месяца в колледже!

Как хорошие исследователи, Дженкинс и Далленбах (Jenkins & Dallenbach, 1924) внимательно отнеслись к вопросу контроля за экспериментами и использовали многие из тех процедур, о которых вы узнали в данном курсе. Например, они использовали списки из 10 бессмысленных слогов (стимул согласный-гласный-согласный) и при этом демонстрировали в отдельный момент времени только один слог. Испытуемые в ходе каждой попытки читали слоги вслух до тех пор, пока не могли их точно воспроизвести (операциональное определение изучения). Чтобы обеспечить одинаковое произнесение слогов двумя участниками, некоторые гласные сверху были помечены чертой, означающей, что звук «долгий»; непомеченные гласные произносились как «краткие». Чтобы избежать систематичности в последовательности кратких гласных (потенциальный осложнитель), «отметки были поставлено беспорядочно, но так, чтобы краткие и долгие гласные в списках в различных позициях встречались одинаково часто» (р. 606). Это значит, что для контроля последовательности долгих и кратких гласных в списках исследователи использовали позиционное уравнивание. Также с помощью позиционного уравнивания они избежали эффекта последовательности в смене интервалов разной длины, предшествующих воспроизведению информации, - «промежутки времени между запоминанием и воспроизведением варьировались случайным образом» (р. 607). Время, в которое проходило изучение списков, не изменялось. При условии «бодрствование» студенты запоминали списки между 8 и 10 часами утра, далее занимались своей обычной студенческой работой, а затем через 1,2,4 или 8 часов возвращались в лабораторию и воспроизводили информацию. При условии «сон» списки изучались между 11:30 вечера и 1 ночи. Затем студенты ложились спать, а через 1,2,4 или 8 часов их будил Карл Далленбах для проверки запоминания. У исследования был один потенциальный осложнитель - при условии бодрствования студентов просили вернуться в лабораторию для воспроизведения информации (т. е. они знали время, прошедшее с момента запоминания), а при попытках со сном они не знали, когда их разбудят. Дженкинс и Далленбах знали об этой проблеме, рассматривали альтернативные варианты, но решили, что их процедура подходит для проведения исследования.

Каковы были результаты? На рис. 8.5 воспроизводится график, приведенный в статье, посвященной этому исследованию. Он отображает данные, полученные для каждого студента. Каждая точка на графике представляет собой среднее значение для восьми (обычно) попыток, сделанных для каждого экспериментального условия. График имеет несколько особенностей. Во-первых, он говорит о том, что студенты выполняли задания сходным образом. Во-вторых (и это было заметным открытием), видно, что в сравнении с ситуацией бодрствования после сна воспроизведение улучшалось. В-третьих, график показывает наличие взаимодействия. По словам Дженкинса и Далленбаха, «кривые экспериментов с бодрствованием имеют хорошо известный вид: резкое падение, а затем кривая постепенно становится более горизонтальной. Однако вид кривых для экспериментов со сном заметно отличается: после небольшого падения кривая изменяет направление и идет горизонтально на довольно высоком уровне» (1924, р. 610).

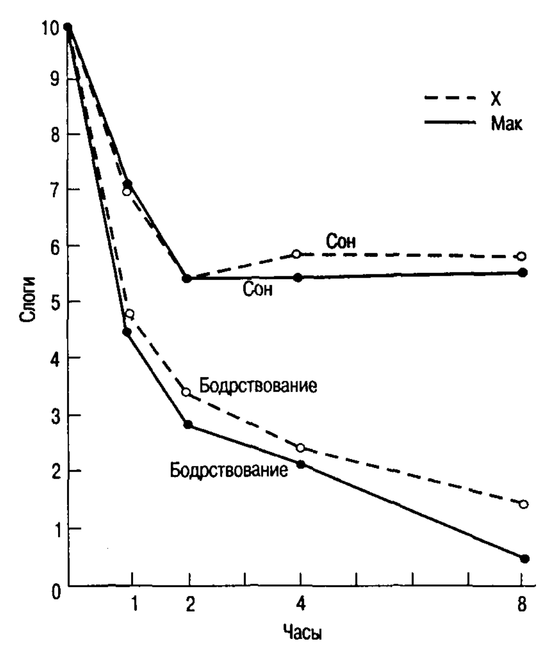


Рис. 8.5. Данные исследования ретроактивного влияния Дженкинса и Далленбаха для обоих испытуемых - студентов Корнельского университета, Л. Р. Ходелла (X) и Дж. С. Мак-Грю (Мак). Имейте в виду, что исследование было проведено задолго до того, как в соответствии с этическим кодексом стало необходимо скрывать имена участников (Jenkins & Dallenbach, 1924)

Исследование показало еще один очень интересный результат, который никогда не упоминают в учебниках. По ходу экспериментов Карлу Далленбаху стало очень трудно будить студентов (насколько трудно было самому Далленбаху просыпаться в разное время среди ночи, нигде не упоминается). Далленбаху также было трудно «определить, проснулись ли они. Наблюдатели вставали, шли в соседнюю комнату, воспроизводили слоги, а на следующее утро они ничего не помнили» (Jenkins & Dallenbach, 1924, p. 607)! В то время полусонное состояние считалось близким к гипнотическому, а поэтому Дженкинс и Далленбах набрали других студентов и частично повторили исследование, но вместо сна разной длительности они просили испытуемых учить и воспроизводить списки слогов, находясь в состоянии гипноза, и при этом использовали различные промежутки между этими событиями. Исследователи обнаружили, что в этом случае воспроизведение информации было практически полным вне зависимости от промежутков времени. Это указывает на явление, впоследствии названное когнитивными психологами «зависимостью научения от состояния».

Виды факторных планов

Аналогично рис. 7.1, на котором показано дерево принятия решений для однофакторных планов, рис. 8.6 представляет решения, которые необходимо принять для получения одного из шести видов факторных планов. Четыре из них повторяют планы, указанные на рис. 7.1, а два возможны только в случае нескольких переменных. Во-первых, тогда как в однофакторном плане независимая переменная может

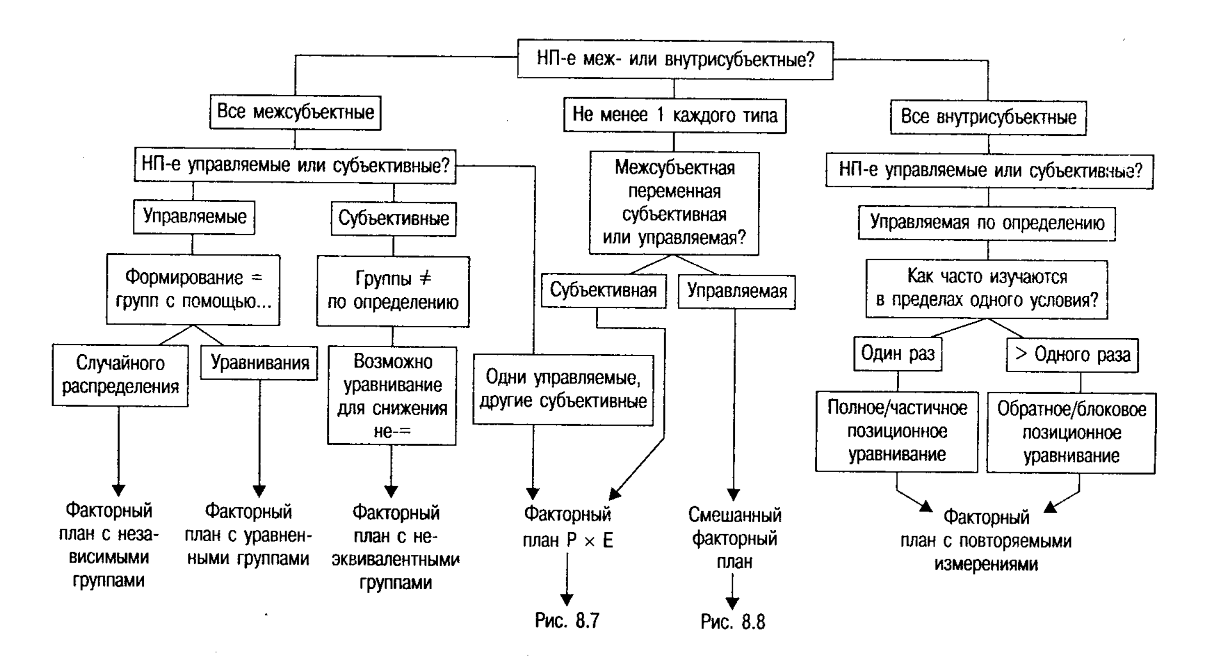


Рис. 8.6. Дерево принятия решений – факторные планы

быть либо межсубъектной, либо внутрисубъектной, факторный план может содержать обе эти переменные. В таком случае план называется **смешанным факторным планом.** В смешанном плане по крайней мере одна переменная должна изучаться как межсубъектная и одна — как внутрисубъектная. Во-вторых, некоторые межсубъектные факторные планы содержат как управляемые независимые переменные, так и субъектные. Поскольку такой план может выявить взаимодействие между типом испытуемого и созданными в эксперименте условиями, его можно назвать **факторным планом Р** х **Е** (person x environment), т. е. «планом "индивидуум—среда"», где «среда» понимается в широком смысле как любая управляемая независимая переменная. Планы Р х Е различаются в зависимости от того, является управляемая переменная межсубъектной или внутрисубъектной. В последнем случае план Р х Е называется смешанным планом Р х Е. На рис. 8.7 и 8.8 показаны решения, которые необходимо принять, чтобы получить смешанный план или план Р х Е. Рассмотрим их более подробно.

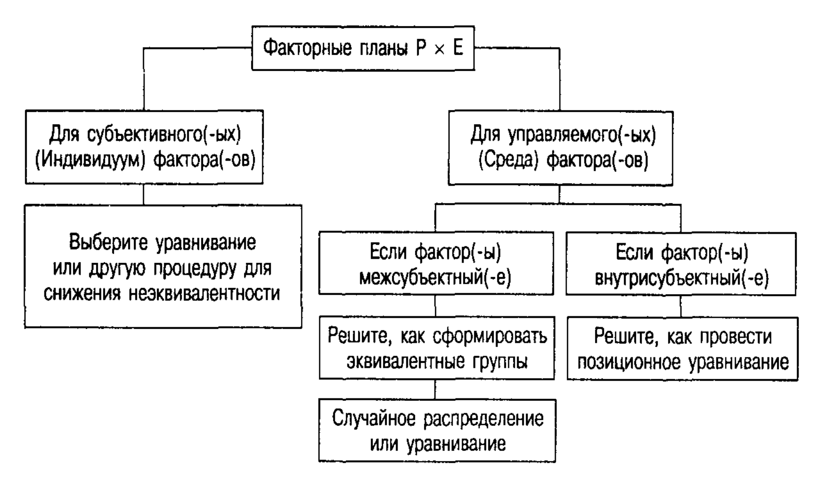


Рис. 8.7. Решения, которые необходимо принять для получения плана Р х Е

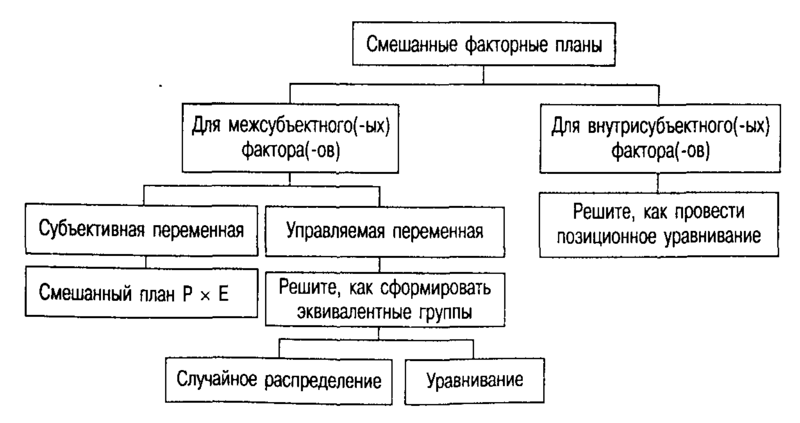


Рис. 8.8. Решения, которые необходимо принять для получения смешанного плана

Смешанные факторные планы

В главе 6 вы узнали, что если независимая переменная является межсубъектной, то создание эквивалентных групп может вызвать проблемы, которые решает случайное распределение. Если независимая переменная внутрисубъектная, трудности связаны с возможным появлением эффекта последовательности — такие проблемы решаются с помощью позиционного уравнивания. При использовании смешанных планов исследователи обычно встречаются и с проблемой эквивалентных групп, и с проблемой эффекта последовательности. Но это происходит не всегда — есть один вид смешанного плана, при котором позиционное уравнивание не используется, поскольку в ходе исследования изучается сам эффект последовательности. Например, при исследовании обучения и запоминания часто используется внутрисубъектная независимая переменная «попытки». В таком случае не имеет смысла проводить позиционное уравнивание, так как одна из целей исследования заключается в том, чтобы продемонстрировать последовательные изменения от попытки к попытке. Следующие два примера иллюстрируют смешанные планы; один из них требует применения позиционного уравнивания, а во втором переменная «попытки» измеряется многократно (план с повторяемыми измерениями).

Пример 16. Смешанный факторный план с позиционным уравниванием

Рискинд и Мэдакс (Riskind & Maddux, 1993), по-видимому, смотрели слишком много плохих фильмов ужасов про пауков, и это навело их на мысль проверить, как люди управляют своими эмоциями в ситуациях, вызывающих страх. Они построили смешанный факторный план 2 χ 2, в котором управляли двумя параметрами: самоуважением и «устрашающим приближением». Самоуважение — это чувство собственной компетентности в решении жизненных проблем. Обычно его изучают как субъективную переменную, но в данном исследовании им управляли экспериментаторы. Участников попросили представить ситуацию, в которой они сидят на стуле в маленькой комнате, по которой бегает тарантул. Участникам, выполняющим задание в условиях высокого самоуважения, велели представить, что дверь в комнату открыта, что они могут свободно передвигаться, а под рукой имеется журнал, которым можно при необходимости ударить паука. Участникам, случайным образом распределенным в группу с условием низкого самоуважения, велели представить, что они крепко привязаны к стулу, дверь закрыта, а журнал лежит вне досягаемости. Пока участники представляли себе подобные обстоятельства, им показывали фильмы про пауков, в которых они а) сидели неподвижно или убегали или б) приближались (переменная «устрашающее приближение»). Действие этой переменной оценивалось многократно, а ее значения изменялись в порядке, полученном с помощью позиционного уравнивания (к сожалению, авторы не указали, какая именно процедура позиционного уравнивания была использована). Зависимой переменной была сила страха, самостоятельно оцениваемая испытуемыми.

Результаты этого исследования, представленные на рис. 8.9 в матричном и графическом виде, хорошо иллюстрируют ситуацию приоритета взаимодействия перед основным эффектом. Несложно заметить различия между значениями среднего арифметического по рядам и столбцам, при этом оба основных эффекта являются статистически значимыми. Но гораздо более важно, что между факторами

было обнаружено взаимодействие. Участники испытывали очень сильный страх (4,50) при создании ситуации фильма ужасов (надвигающиеся пауки плюс низкое самоуважение), а при трех других условиях они сообщили о среднем или низком уровне страха (2,24, 2,64, 2,73). Таким образом, у участников с высоким самоуважением уровень страха был невысоким вне зависимости от движения пауков, тогда как у испытуемых с низким самоуважением он зависел он того, приближаются или удаляются пауки.

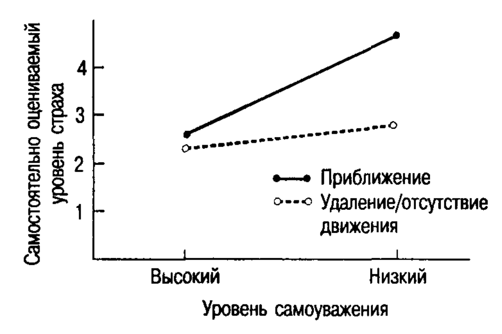


Рис. 8.9. Влияние «угрожающего приближения» и самоуважения на чувство страха (Riskind & Maddux, 1993)

Пример 17. Смешанный факторный план без позиционного уравнивания

Хороший пример ситуации, когда переменная «попытки» является внутрисубъектной при смешанном экспериментальном плане, — это процедура исследования памяти, называемая «падение проактивного влияния» (Wickens, Born, & Allen, 1963). Проактивное влияние — это явление, при котором запоминанию и воспроизведению новой информации препятствует предшествующее им запоминание старой информации. Вы могли встречаться с таким явлением, если вам не удавалось запомнить новый номер телефона из-за того, что старый постоянно приходил в голову. Предполагалось, что если новая информация сходна со старой, то влияние будет особенно сильным. Один из способов проверки идеи о том, что сила проактивного влияния зависит от сходства элементов, состоит в том, чтобы построить эксперимент, в котором испытуемые будут запоминать и воспроизводить последовательность похожих стимулов, а затем перейдут к изучению другого вида стимулов. Предположительно, в случае похожих стимулов проактивное влияние должно нарастать от попытки к попытке, а затем «упасть» при смене стимулов. Это означает, что точность воспроизведения информации при нарастании проактивного влияния должна постепенно снижаться, а затем снова повыситься при его падении.

При изучении падения проактивного влияния в качестве стимулов обычно используются слова или бессмысленные слоги, но в исследовании Гантера, Берри и Клиффорда (Gunter, Berry, & Clifford, 1981) был применен более реалистичный подход: в своих экспериментах исследователи использовали фрагменты телевизионных новостей. Мы рассмотрим только первый эксперимент, отлично иллюстрирующий смешанный факторный план без позиционного уравнивания.

Участникам сказали, что им покажут фрагменты телевизионных новостей, а затем проверят, насколько хорошо они запомнили содержание показанного материала. В ходе каждой попытки испытуемые просматривали по три истории, затем выполняли короткое отвлекающее задание (кроссворд), после чего старались как можно подробнее воспроизвести увиденное. Каждый участник выполнял по четыре таких попытки. Половина испытуемых была случайным образом распределена в группу с условием падения проактивного влияния; они выполняли три попытки, в которых все новости были посвящены одной теме (внутренние политические события), а затем четвертую («падение влияния»), с новостями на другую тему (международные политические события). Вторая половина испытуемых вошла в контрольную группу; для них новости во всех четырех попытках были на одну тему. Таким образом, использовался смешанный факторный план 2 (падение/отсутствие падения) x 4 (попытки). Межсубъектным фактором было, встречались или нет участники с условием падения влияния. Попытки рассматривались как повторно измеряемый или внутрисубъектный фактор.

На рис. 8.10 показаны результаты этого исследования. Наличие проактивного влияния видно из оценок, полученных контрольной группой, которые постепенно снижаются. Оценки, полученные экспериментальной группой, говорят о падении проактивного влияния.

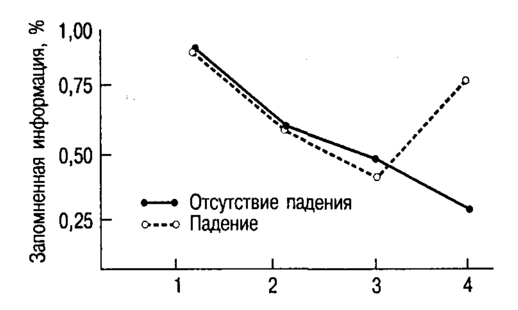


Рис. 8.10. Падение проактивного влияния (Gunteretal., 1981)

Следует отметить одну особенность контроля за исследованием. Возможная проблема заключалась в том, что улучшение выполнения задания в ходе попыток с падением влияния могло быть связано с тем, что международные новости запомнить легче, чем новости внутренней политики. Чтобы исключить такую возможность, половина испытуемых в экспериментальной группе в ходе трех попыток смотрела внутренние политические новости, а в ходе четвертой — международные, тогда как вторая половина смотрела сначала международные новости, а затем внутренние. Аналогично одна половина контрольной группы в ходе четырех попыток смотрела внутренние новости, а вторая — международные. Порядок следования новостей не оказал никакого влияния на результаты[1].

Обозначив план как 2x2, мы выделили две главные переменные, но фактически это был факторный план 2x2x4, где второе «2» означает межсубъектный фактор категории новостей — международные политические новости или внутренние.

Факторные планы с субъективными и управляемыми переменными: планы P x E

В главе 5 вы познакомились с понятием субъективной переменной, которая представляет собой некоторую уже существующую характеристику испытуемого, например возраст, пол или особенность личности. Также вы узнали, что необходимо очень осторожно относиться к выводам при наличии в исследовании субъективных переменных. При условии тщательного контроля за экспериментом выводы о существовании причинно-следственной связи можно делать только при использовании управляемых переменных, в случае субъективных переменных выводы сделать невозможно. Факторные планы P x E включают как субъективные (P — *person,* индивидуум), так и управляемые (E — *environment,* среда) переменные. В таком случае при наличии значимого основного эффекта у управляемого фактора «среда» выводы о существовании причинно-следственной связи делать можно, а если основной эффект обнаружен у субъективной переменной, или «личностного» фактора, то нельзя. Также невозможно делать выводы при наличии взаимодействия. Несмотря на эти ограничения, планы, в которых используются и субъективные, и управляемые переменные, весьма распространены, отчасти из-за того, что они объединяют два исследовательских направления, описанных Вудвортсом в его известной колумбийской библии (см. главу 5). Корреляционное направление связано с изучением индивидуальных различий, а субъективная переменная, или фактор P в плане P x E, призвана обнаружить эти различия. Наличие значимого основного эффекта у этого фактора означает, что два различных типа людей действуют по-разному, причем их поведение измеряется как зависимая переменная. Экспериментальное направление посвящено исследованию общих законов поведения, которые в определенной степени можно применить ко всем людям вне зависимости от индивидуальных различий. В такой ситуации обнаружение значимого основного эффекта у управляемой переменной, или фактора E в плане P x E, показывает, что ситуативный фактор достаточно силен, чтобы повлиять на поведение различных типов людей. Рассмотрим гипотетический пример, в котором сравниваются интроверты и экстраверты (переменная P). В ходе исследования испытуемые решают задачи либо в небольшой заполненной людьми комнате, либо в большой и свободной (переменная E). Допустим, что получены следующие результаты (ЗП = количество решенных проблем):

	Маленькая комната	Большая комната
Интроверты	12	12
Экстраверты	18	18

В данном случае наблюдается основной эффект у типа личности, отсутствие основного эффекта у окружения и отсутствие взаимодействия. Очевидно, что экстраверты выполняют задание лучше, чем интроверты (18 > 12) вне зависимости от условий. В такой ситуации исследователь мог бы найти важные различия между этими типами личности и обобщить их на другие ситуации.

Из следующих результатов можно сделать совершенно другой вывод:

	Маленькая комната	Большая комната
Интроверты	12	18
Экстраверты	12	18

Это указывает на основной эффект фактора среды, а также на отсутствие основного эффекта для типа личности и отсутствие взаимодействия. В данном случае сильно влияние среды (размер комнаты) (18 > 12) и это влияние не ограничивается одним типом личности; вне зависимости от типа личности в условиях скопления людей задание выполняется хуже. Таким образом, обнаружение значимого основного эффекта для фактора Р свидетельствует о наличии сильных личностных различий, тогда как значимый основной эффект для фактора Е говорит о силе воздействия окружающей обстановки, выходящего за пределы одного типа личности. Кроме того, в результате такого исследования можно получить два основных эффекта, показывающих, что важны оба фактора.

Но самым интересным результатом, который можно получить при использовании плана Р х Е, является взаимодействие. Наличие взаимодействия означает, что на один тип людей изменение окружающей обстановки влияет одним образом, а на другой — другим. Предположим, что в том же примере с интровертами и экстравертами получены следующие результаты:

	Маленькая комната	Большая комната
Интроверты	18	12
Экстраверты	12	18

В таком случае ни один из основных эффектов не будет значимым, но обнаружится взаимодействие. В случае интровертов события развивались одним образом, а в случае экстравертов — другим: интроверты выполняли задание намного лучше в маленькой комнате, а экстраверты — в большой.

Факторные планы, включающие и субъективные, и управляемые переменные, часто используются в исследованиях, посвященных обучению или проверке эффективности психотерапии (Smith & Sechrest, 1991). О важности обнаружения значимых взаимодействий в таких планах свидетельствует то, что их также называют **ATI-планами,** или «планами взаимодействия способностей и внешнего воздействия» *{Aptitude-TreatmentInteraction).* Как вы, возможно, догадываетесь, «способности» представляют собой неуправляемую субъективную (личность) переменную, а «воздействие» — управляемую переменную среды. Примером исследования эффективности психотерапии является работа Абрамовича с соавторами (Abramovitz, Abramovitz, Roback, & Jackson, 1974). Переменной Р был локус контроля. Люди с внешним локусом контроля считают, что их жизнь зависит от внешних обстоятельств, а имеющие внутренний локус контроля уверены в том, что события их жизни являются следствием их собственных действий и решений. В ходе исследования испытуемые с внешним локусом контроля показали лучшие результаты при директивной терапии, задающей направление их деятельности, чем при недирективной, возлагающей на пациента больше ответственности за прогресс. Результаты участников с внутренней точкой контроля были обратными:

прогресс был весьма заметным в случае косвенной терапии и довольно слабым в случае прямой.

При исследовании процесса обучения ATI-планы обычно используют, если фактор способностей, или личностный фактор, — это переменная «тип обучения», а фактор воздействия, или среды, — инструкция, даваемая участникам. Например, на рис. 8.11 показаны результаты исследования обучения, проведенного Валери Шут — крупным авторитетом в области ATI-планов (Shute, 1994). В ходе исследования сравнивались два подхода к обучению основным законам электричества — формирование правила и применение правила. Участники были случайным образом распределены на две группы, соответствующие двум подходам. Субъективная переменная имела два значения: высокая и низкая оценка «исследовательского» поведения. График показывает, что испытуемые, получившие высокую оценку исследовательского поведения, лучше справились с обучением по методу формирования правила, в ходе которого их просили больше работать самостоятельно, а получившие низкую оценку показали более высокие результаты обучения по методу применения правила, которое было более строго организовано.

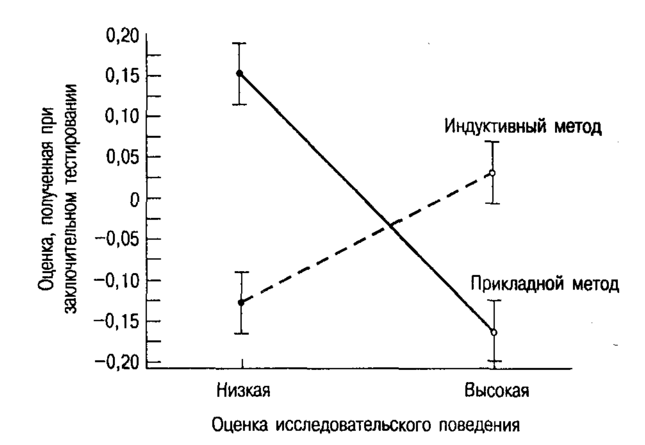


Рис. 8.11. P x E взаимодействие между уровнем исследовательского поведения и методом обучения (Shute, 1994)

Также факторные планы P x E весьма популярны в исследованиях личности и в психологии здоровья. Приведенный ниже пример отлично иллюстрирует исследование из области психологии здоровья, в котором рассматривается личностный фактор.

Пример 18. Факторный план с взаимодействием P x E

В последние 25 лет был заметен значительный интерес к особенностям личности людей, обладающих так называемым поведением типа А (Friedman & Rosenman, 1974). Индивидуумы типа А любят соревноваться, нацелены на успех и навязчиво внимательны ко многим вещам, включая время (они всегда приходят раньше и никогда не опаздывают). Они стараются одновременно делать много дел и выглядят людьми, обладающими очень большим запасом энергии. К сожалению, люди типа А, сочетающие эти черты с недружелюбным характером, склонны к развитию коронарной болезни. Это означает, что поведение типа А в некоторых обстоятельствах может быть весьма нездоровым.

Исследования поведения типа А часто строятся в соответствии с факторным планом P x E. Переменной P (субъективной) является тип поведения — тип А или его противоположность, более спокойный тип В. Отбор участников для двух групп проводится в соответствии с результатами теста на поведение типа А/В. Переменная E — это некоторый фактор, которым управляет экспериментатор. К примеру, в исследовании Холмса с соавторами (Holmes, McGilley, & Houston, 1984) принадлежность студентов колледжа к типу А или В определялась по результатам студенческой версии «Исследования активности» Дженкинса — теста, широко применяемого при исследовании поведения типа А. Было протестировано 394 студента, и из всех оценок исследователи выбрали 30 наибольших и 30 наименьших. Участников в двух группах случайным образом распределили по трем заданиям, различающимся по степени трудности. Таким образом, управляемой переменной была трудность задания. Заданием являлся арифметический тест, позаимствованный из общего IQ-теста. Участники прослушивали последовательности чисел (например, 3-4-8) и должны были повторить их в обратном порядке (8-4-3). Три уровня сложности различались по количеству чисел, предлагаемых каждому участнику в ходе шести попыток: два, пять или семь. Зависимой переменной было возбуждение, проявляющееся в процессе выполнения задания. Оно операционально определялось в терминах нескольких физиологических характеристик, включая систолическое кровяное давление.

На рис. 8.12 показано взаимодействие P x E. Возбуждение у студентов типа А и типа В не различалось в случае легкого и умеренно сложного задания, но при очень сложном задании систолическое давление у типа А продолжало подниматься, тогда как у типа В оно оставалось неизменным. Таким образом, сравнение типов А и В показало значимое увеличение систолического кровяного давления, но только в случае действительно сложного задания. Было обнаружено отсутствие основного эффекта для типа личности и его наличие для трудности задания, как видно из общего увеличения кровяного давления, происходящего по мере нарастания сложности работы. В целом, исследование показало, что «разница возбуждения... между людьми, принадлежащими к типам А и В, проявляется при высоком уровне сложности» (Holmes et al., 1984, p. 1326). Это открытие согласуется с соревновательной, нацеленной на успех природой людей, принадлежащих к типу А.

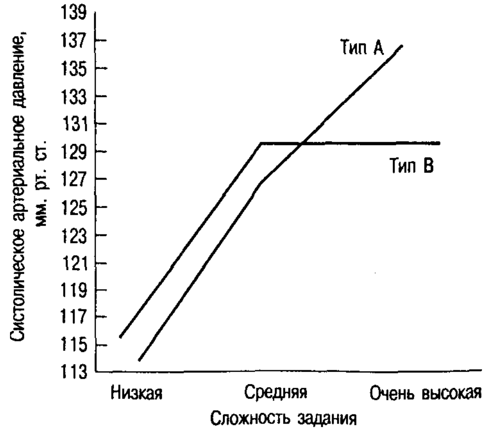


Рис. 8.12. P x E-взаимодействие: Типы личности А и В и уровень сложности задания (Holmes et al., 1984)

Кстати, защитники использования планов P x E уважают работы Курта Левина (1890-1947), пионера социальной и детской психологии. Лейтмотивом работ Левина была мысль о том, что для полного понимания поведения необходимо изучать как свойства личности, так и окружение, в котором действует человек. Ученый выразил эту идею с помощью хорошо известной формулы: B =*f*(P, E), поведение *(behavior)* — это функция личности *{person)* и среды *(environment)* (Goodwin, 1999). Факторные планы P x E, унаследовавшие свое название от формулы Курта Левина, идеально подходят для исследования видов взаимоотношений, которые, по мнению Левина, характеризуют человеческое поведение'.

Набор участников для факторных исследований

Из определений различных факторных планов должно быть ясно, что в зависимости от плана число участников, требующихся для проведения исследования, может заметно различаться. К примеру, если вам, чтобы заполнить одну ячейку матрицы факторного плана 2x2, нужно 5 участников, то общее число испытуемых, которых потребуется набрать, будет равнятся 5, 10 или 20. Рисунок 8.13 объясняет, почему это так. На рис. 8.13, а обе переменные исследуются как межсубъектные, а следовательно, на каждую ячейку потребуется 5 участников, итого 20. На рис. 8.13, б обе переменные являются внутрисубъектными, а план — факторным с повторяемыми измерениями. Данные для всех четырех ячеек будут получены от одних и те же пятерых участников. В случае смешанного плана (рис. 8.13, в) одна переменная межсубъектная и одна — внутрисубъектная, поэтому данные для двух ячеек будут собраны с первых пяти участников, а двух других — со вторых. Общее число участников будет равняться 10.

За определением количества испытуемых, которых необходимо набрать для проведения экспериментов, следует вопрос о том, как обращаться с участниками исследования. Вставка 8.2 содержит практическое руководство по этике исследований.

ВСТАВКА 8.2

Этика — компетентность и этичность исследователя

В главе 2 вы узнали о существовании этического кодекса *APA,* а во всех последующих главах встречались со вставками, посвященными этике исследований. Хотя у вас должны были сложиться весьма четкие представления об этических требованиях к проведению исследований (согласие, дебрифинг и т. д.), могло остаться непонятным, как все это применять на практике. Настал момент представить вам набор практических советов о том, как быть действительно ответственным экспериментатором.

Недостаточно являться на место проведения экспериментов к моменту их начала - приходите заранее, чтобы успеть подготовить все необходимые материалы до прихода участников.
Обращайтесь с участниками исследования всегда так вежливо и уважительно, как вы бы хотели, чтобы они обращались с вами при обратном распределении ролей. Приветствуйте их, когда они приходят в лабораторию, и благодарите за то, что они записались на проведение экспе-

' Одно неудачное следствие из использованного Левином обозначения P *{person)* состоит в том, что план P x E подразумевает участие в исследовании только люден. Тем не менее подобные исследования довольно часто проводятся с участием животных (например, исследования, в которых субъективной переменной является пол изучаемых приматов).

риментов и пришли. Они могут быть слегка напуганы тем, что может случиться с ними во время «психологического» эксперимента, поэтому ваша первая задача - помочь им расслабиться и в то же время показать себя как профессионала, ответственного за ход исследования. Никогда не забывайте, что участники делают вам одолжение - обратное неверно. Почаще улыбайтесь.
Начните работу с получения обоснованного согласия и старайтесь не создавать у участников впечатления, что это - отнимающая много времени формальность, которую необходимо исполнить, прежде чем начать важную работу. Дайте участникам понять, что вы хотите, чтобы они ясно понимали свои будущие действия. Если они не задают вопросов во время чтения документа о согласии, по окончании чтения спросите, имеются ли у них вопросы. Проследите, чтобы участники подписали две копии документа о выражении согласия - одну для себя и вторую, необходимую для составления документации по исследованию.
Неплохо заранее написать «инструкцию для участников». В зависимости от исследования можно прочитать инструкцию участникам или раздать им распечатанные копии. Не обязательно детально разрабатывать письменную инструкцию, можно составить перечень основных моментов процедуры исследования.

Прежде чем проводить «реальный» эксперимент, отработайте роль экспериментатора с друзьями или коллегами по лаборатории. Проиграйте весь ход эксперимента - представьте, что это генеральная репетиция, в ходе которой вы сможете исправить возникшие проблемы.
Будьте внимательны к любым признакам стресса у участников. Это может означать изменение хода исследования и отказ от данных, полученных от испытуемого, но благополучие участников важнее, чем собранные вами данные. Кроме того, вы не профессиональный консультант, а поэтому, если участники испытывают беспокойство из-за участия в исследовании, вежливо препроводите их к своему руководителю.

Тщательно подготовьте дебрифинг. Вам как студентам-экспериментаторам скорее всего, не придется проводить исследования с серьезной мистификацией или высоким уровнем стресса, но вы должны сделать так, чтобы опыт, полученный вашими испытуемыми, имел для них образовательное значение. Поэтому вам необходимо упростить описание явлений, которые вы надеетесь обнаружить в ходе исследования, а также дать участникам возможность высказать свои пожелания по поводу экспериментальной процедуры или идеи о проведении будущих исследований. Не торопитесь заканчивать дебрифинг и не старайтесь дать лишь краткое описание событий в надежде, что испытуемые быстро уйдут. А если они захотят уйти, не пройдя дебрифинг (такое иногда бывает), не разрешайте им. Дебрифинг - это важная часть вашей работы, за которую вы несете ответственность как исследователь и человек, передающий свои знания другим людям. (Конечно, если вам скажут «Я думал, вы сказали, что мы можем уйти в любой момент», то вы уже ничего не сможете поделать!)
Прежде чем испытуемые уйдут, напомните им, что на имеющихся у них документах о согласии есть информация о людях, с которыми можно связаться, если возникнут вопросы по поводу исследования. Проинформируйте участников, когда планируется закончить исследование и когда можно будет узнать его результаты. Также попросите их не обсуждать эксперимент с другими людьми, которые могут оказаться его участниками. **Утечка,** или склонность испытуемых рассказывать об исследовании, в котором они принимали участие, может вызвать серьезные проблемы, особенно в небольших учебных заведениях (более подробно об обязанностях испытуемых см. вставку 6.3). Если вы хорошо обращались с участниками в ходе исследования, то вероятность такого общения между ними возрастает.
Перед уходом участников поблагодарите их за то, что они потратили на исследование свое время и силы, и не забудьте улыбнуться им на прощание. Помните о том, что некоторые студенты, которых вы тестируете, еще не выбрали направление специализации и, возможно, думают о психологии - участие в вашем исследовании может усилить их интерес.

а) Для плана 2 χ 2 с четырьмя различными группами и пятью участниками на каждую ячейку требуется 20 испытуемых

S1 S2 S3 S4 S5	S11 S12 S13 S14 S15
S6 S7 S8 S9 S10	S16 S17 S18 S19 S20

б) Для плана 2 χ **2** с повторяемыми измерениями и пятью **участниками на** каждую ячейку требуется 5 испытуемых

S1 S2 S3 S4 S5	S1 S2 S3 S4 S5
S1 S2 S3 S4 S5	S1 S2 S3 S4 S5

в) Для смешанного плана 2 χ **2** с пятью участниками на каждую ячейку **требуется 10 испытуемых**

S1 S2 S3 S4 S5	S1 S2 S3 S4 S5
S6 S7 S8 S9 S10	S6 S7 S8 S9 . SO

Рис. 8.13. Количество участников в факторных **планах**

Анализ факторных планов

Как мы уже знаем, многоуровневые однофакторные планы анализируются с помощью однофакторного метода *ANOVA.* Аналогично факторные планы, в которых данные получают с помощью интервальной шкалы или шкалы отношений, анализируются Л^-факторным методом *ANOVA,* где ^означает количество используемых независимых переменных. Так, факторный план 2x3 будет анализироваться с помощью двухфакторного дисперсионного анализа, а 2 x 2 x 4 — трехфакторного.

В процессе однофакторного дисперсионного анализа вычисляется только одно отношение *F,* а далее, если /"значимо, может проводиться дополнительная проверка. В случае факторного плана вычисляется более одного отношения *F.* Для каждого основного эффекта или взаимодействия находится свое *F.* Например, для плана 2 x 2, построенного для изучения влияния визуальных образов и скорости показа на память, одно отношение F будет находиться, чтобы определить наличие основного эффекта у переменной «способ запоминания», второе — чтобы обнаружить основной эффект для скорости показа, а третий, чтобы выявить возможное взаимодействие между этими двумя переменными (в приложении *С* приведен пример дисперсионного анализа 2 x2). В случае факторного плана А xB x С вычисляется *семь* отношений: три для основных эффектов у *A, B* и *C,* еще три для эффектов двухфакторного взаимодействия A x B, B x C и AC плюс один для трехфакторного взаимодействия A x B x C. Так же как и в случае с однофакторным методом *AN O VA,* факторный дисперсионный анализ может включать последующую проверку. Например, получение в ходе анализа 2 x 3 значимого /для фактора, принимающего три значения, приведет к дополнительной проверке (например, к вычислению *HSD* Таки — см. приложение *С),* в ходе которой будут сравниваться результаты для уровней 1 и 2, 1 и 3, а также 2 и 3.

Прежде чем завершить эту главу, я хочу сделать еще одно замечание, касающееся факторных планов и дисперсионного анализа. В данной главе мы рассматривали множество матриц, которые могли навести вас на мысль о сельском хозяйстве. Прочитав вставку 8.3, вы узнаете о сэре Рональде Фишере, который изобрел дисперсионный анализ, и поймете, что подобные ассоциации не случайны.

ВСТАВКА 8.3

История — факторные планы на ферме

Представьте, что вы летите в небольшом самолете над Канзасом. В окно вы видите тянущиеся миля за милей поля ферм, расчерченные на прямоугольники. Это может напомнить вам о факторных матрицах 2 x 2 и 3 x 3 , с которыми вы познакомились в этой главе. Это связано с тем, что факторные планы и метод *ANOVA,* применяющийся для их анализа, были разработаны сэром Рональдом Фишером при проведении сельскохозяйственного исследования. Эмпирический вопрос был: «Каковы наилучшие условия или комбинации условий для выращивания урожая X?»

Рональд Эйлмер Фишер (1890-1962) был одним из известнейших ученых, занимавшихся статистикой, равный по заслугам Карлу Пирсону, открывшему способ измерения корреляции, сегодня обозначаемой как «Пирсонова *r*» (см. следующую главу). Фишер создал статистическую процедуру, использующуюся для проверки предположений в области генетики,

но среди психологов он известен созданием дисперсионного анализа, в ходе которого находится отношение F, позволяющее сделать вывод о нулевой гипотезе экспериментального исследования. Несложно догадаться, что означает «F»[1].

Начиная с 1920 г. на протяжении около 15 лет Фишер работал на экспериментальной сельскохозяйственной станции в английском городе Ротамстед. Он участвовал в исследовании воздействия на урожай различных переменных, например вида удобрений, уровня осадков, времени посадки, а также генетических особенностей растений. Фишер опубликовал статью «Исследование сельскохозяйственных культур. VI. Эксперименты по изучению реакции картофеля на углекислый калий и азот» (Kendall, 1970, p. 447). В ходе исследований он открыл дисперсионный анализ и использовал его для анализа полученных данных. Особенно он подчеркивал важность использования факторных планов «так как при раздельных (однофакторных) экспериментах мы никогда не обнаружим *взаимодействий* между различными компонентами» (Fisher, 1935/1951, p. 95, курсив мой. - *Дж. Г.).* Сельскохозяйственные урожаи зависят от сложных комбинаций факторов, и их раздельное изучение не позволяет оценить взаимодействие между ними.

Простой план 2 x 2 одного из экспериментов Фишера изображен на рис. 8.14. Каждый прямоугольник показывает, как обрабатывался один из небольших участков земли. Как и любой факторный план, этот план позволяет оценить основные эффекты (в данном случае для вида удобрений и сорта пшеницы), а также взаимодействие этих факторов. Если предположить, что закрашенное поле производит значительно больше пшеницы, чем остальные три (их урожаи одинаковы), то можно обнаружить наличие взаимодействия: удобрения работали эффективно только для определенного сорта пшеницы.

	Экспериментальное удобрение	Отсутствие экспериментального удобрения
Пшеница: генетическая особенность I	Пшеничное поле *A*	Пшеничное поле *B*
Пшеница: генетическая особенность II	Пшеничное поле C	Пшеничное поле D

Рис. 8.14 Взаимодействие в сельском хозяйстве

Фишер включил свою работу по дисперсионному анализу в труд по статистике, опубликованный им в 1925 г. (через год после того, как Дженкинс и Дапленбах издали исследование взаимосвязи сна и памяти) (Fisher, 1925). Самая известная из его работ, «Разработка экспериментов» (Fisher, 1935/1951), увидела свет через 10 лет после этого. Она посвящена методу ***ANOVA*** и включает обсуждение вопросов статистики и методов исследований. Дисперсионный анализ и факторные планы далеко не сразу получили широкую известность в Соединенных Штатах, но с начала 50-х гг. психологи-экспериментаторы стали использовать их в качестве основного метода статистического анализа.

На этом мы завершаем рассказ об экспериментальных планах, которому были посвящены две последние главы. Прежде чем вы научитесь свободно говорить на

[1] От англ. *farm* - ферма.

«языке экспериментальной психологии» и создавать методологически правильные эксперименты для проверки ваших гипотез, вам потребуется несколько раз прочитать эти главы (а также главы 5 и 6) и попрактиковаться в использовании различных экспериментальных планов. Далее мы приступим к рассмотрению другого направления исследований, посвященного не изучению различий, а выявлению степени сходства изучаемых переменных.

Резюме

Основы факторных планов

Факторные планы используются для изучения действия нескольких независимых переменных. Они обозначаются с помощью системы обозначений, указывающей количество независимых переменных, значений, принимаемых каждой из переменных, и экспериментальных условий. Например, факторный план 2 x 3 («два на три») включает две независимые переменные, первая из которых принимает два значения, а вторая — три, и шесть различных условий (2 умножить на 3).

Результаты: основной эффект и взаимодействие

Общее влияние независимой переменной в факторном плане называется основным эффектом. В плане 2 x 3 могут существовать два основных эффекта: один для переменной с двумя уровнями, а второй для переменной с тремя уровнями. Основное преимущество факторных планов перед исследованиями с одной независимой переменной состоит в том, что они учитывают возможность существования взаимодействий между факторами. При наличии взаимодействия влияние одной независимой переменной различается для разных уровней другой независимой переменной. Результатом факторного исследования могут быть значимые основные эффекты, взаимодействия и то и другое вместе или отсутствие и того и другого.

Виды факторных планов

В факторном плане независимые переменные могут все быть межсубъектными или все внутрисубъектными. Межсубъектные факторные планы могут включать независимые, уравненные или неэквивалентные группы. Внутрисубъектные факторные планы также называются факторными планами с повторяемыми измерениями. Смешанные факторные планы содержат по крайней мере по одному фактору каждого типа (меж- и внутрисубъектному). Факторные планы, содержащие хотя бы одну управляемую и одну субъектную переменную, позволяют обнаружить взаимодействия типа индивидуум x среда (P x E — Person x Environment). Такие взаимодействия показывают, что стимульная ситуация влияет на один тип людей одним образом, а на другой — другим. Основной эффект у фактора P (т. е. у субъективной переменной) отражает существенные различия между типами людей, проявляющиеся в разных условиях. Основной эффект у фактора E (т. е. у управляемой переменной) показывает действие важных внешних факторов на разные типы людей. В исследованиях, посвященных обучению или проверке эффективности психотерапии, взаимодействие между человеком и окружением иногда называется «планами взаимодействия способностей и внешнего воздействия» (ATI-планами).

Задания для повторения

Выбор ответа

1. Исследователь предполагает, что интроверты лучше справятся с решением задач в одиночестве, чем перед лицом публики, а экстраверты, наоборот, выполнят задание лучше в присутствии зрителей. Исследователь:
   1) предполагает наличие основного эффекта у переменной «публика»;
   2) использует смешанный факторный план;
   3) надеется обнаружить эффект взаимодействия;
   4) использует две управляемые переменные.
2. В смешанном факторном плане 3x3 для ячейки A1B1 требуется пять испытуемых. Сколько людей потребуется набрать для этого исследования?
   1) 5;
   2) 10;
   3) 45;
   4) 15.
3. Факторный план 2x3x5 содержит:
   1) 10 различных независимых переменных;
   2) 25 различных условий;
   3) 3 различных независимых переменных;
   4) максимум 30 основных эффектов.
4. Смешанный факторный план 2x2 всегда содержит:
   1) две субъективные переменные;
   2) межсубъектный и внутрисубъектный факторы;
   3) внутрисубъектный фактор и многократно измеряемую переменную;
   4) одну управляемую и одну субъективную переменную.
5. В эксперименте с лабиринтом 30 крыс выполняют задание при включенном свете и 30 — при выключенном. 10 крыс в каждой группе получают подкрепление сразу после выполнения задания, 10 — спустя 5 секунд после выполнения и 10 — через 10 секунд. Что можно сказать о таком исследовании?
   1) обе независимые переменные являются субъективными;
   2) используется план 2 x 2 с независимыми группами;
   3) используется смешанный план 2x3;
   4) исследуются шесть различных условий.

Короткие эссе

1. Чем отличаются понятия «уровней» и «условий» в случае факторных планов?
2. Что такое основной эффект? В терминах содержимого факторной матрицы опишите, как определить наличие основного эффекта.

3. На примере исследования Годдена и Бэддели (шотландский клуб ныряльщиков) проиллюстрируйте факт, что в ходе исследования могут быть получены важные результаты даже при отсутствии основных эффектов.
4. Объясните, почему если в ходе исследования обнаружены и основные эффекты, и взаимодействие, взаимодействие должно рассматриваться первым, а статистически значимые основные эффекты будут иметь гораздо меньше значения для общих результатов исследования.
5. Что такое эффект потолка? Каким образом его можно устранить?
6. Чем отличается смешанный факторный план от плана P x E?
7. На примере исследования с интровертами/экстравертами и размером комнаты покажите, как с помощью плана P x E можно обнаружить, когда а) люди существенно различаются и б) ситуация имеет больше значения, чем индивидуальные различия.
8. Смешанные факторные планы могут включать, а могут не включать позиционное уравнивание. Объясните, почему.
9. Опишите экспериментальный план и общие результаты исследования Дженкинса и Далленбаха, посвященного сну и памяти.
10. Что такое утечка и как ее предотвратить?

Упражнения

Упражнение 8.1. Определения вида экспериментального плана

Для каждого из приведенных ниже исследований определите зависимые и независимые переменные, уровни независимых переменных и вид каждой независимой переменной (межсубъектная или внутрисубъектная; управляемая или субъективная). Определите шкалу измерений для каждой зависимой переменной. С помощью факторной системы нумерации определите количество независимых переменных и их уровней (например, 2 x 3), а с помощью рис. 8.6 — экспериментальный план.

1. На основании оценок «Исследования активности» Дженкинса было создано три группы испытуемых: типа А, типа В и промежуточного типа. Группы состоят из одинакового количества участников, каждый из которых выполняет по два задания. В ходе первого задания испытуемые сидят в небольшой комнате и, не имея часов, должны оценить, когда пройдет две минуты. Второе задание включает такую же оценку, но при его выполнении участники играют в портативные видеоигры.
2. Студентов колледжа в ходе исследования когнитивных карт попросили с помощью прибора для определения направления точно указать направление расположения трех объектов, находящихся на разных расстояниях от лаборатории и не видимых из нее. Один объект — это соседний кампус, второй — близлежащий город, а третий — удаленный город. Половина участников выполняют задание в комнате без окон, в которой находится компас, указывающий на север. Остальные участники выполняют задание в той же комнате, но без компаса.

3. В исследовании тактильной чувствительности у двух равных по количеству групп слепых и зрячих взрослых измеряется порог чувствительности в 10 различных точках тела. Для характеристики порога используются две оценки. Одна половина участников выполняют задание утром, а вторая — вечером.
4. Три группы дошкольников участвуют в исследовании с задержкой вознаграждения, в котором варьируется длительность задержки. Дети во всех трех группах собирают головоломки. Одной группе сказали, что в качестве вознаграждения они могут получить доллар сейчас или три доллара завтра. Вторая группа выбирала между долларом сразу и тремя долларами через два дня, а третья — между долларом сразу и тремя долларами через три дня. В каждой группе половина детей собирала простые головоломки, а вторая половина — сложные. Группы сформированы таким образом, что средний доход родителей одинаков для детей в каждой группе.
5. В исследовании зрительных иллюзий и восприятия размера испытуемые поворачивают ручку, которая регулирует величину одного из двух стимулов. Задача состоит в том, чтобы уравнять стимулы по величине. Каждый участник выполняет по 40 попыток, в ходе каждой из которых измеряется величина ошибки. В одной половине попыток в качестве стимула используются цветные изображения, а во второй — черно-белые. Половину и цветных, и черно-белых изображений демонстрируют с расстояния 10 футов, а другую половину — с 20 футов от участников.
6. В исследовании осознанности чтения студенты шестого курса читают короткий рассказ о бейсболе. Студентов разделили на две группы в зависимости от их познаний в этой игре. В каждой группе одна половина студентов имеет высокие оценки теста на вербальный интеллект, а вторая — низкие.

Упражнение 8.2. Основные эффекты и взаимодействия

Для каждого из приведенных ниже исследований:

1) определите независимые переменные, значения, принимаемые каждой из них, и зависимые переменные;
2) поместите данные в правильные ячейки факторной матрицы;
3) определите наличие основных эффектов и/или взаимодействий;
4) дайте устное описание результатов исследования;
5) постройте график, отображающий результаты.

В целях упражнения предположим, что разница в 2 единицы между значениями среднего арифметического для любых рядов, столбцов или ячеек значима.

1. Исследователь хочет выяснить влияние неопределенности ситуации и количества наблюдателей на оказание помощи. Испытуемые заполняют опросники в пустой комнате или в комнате с двумя людьми, которые выглядят как участники исследования, но таковыми не являются. Экспериментатор раздает опросники, а затем переходит в соседнюю комнату. Через 5 минут раздается громкий удар, похожий на звук падения человека. Половине участников экспериментатор недвусмысленно сообщает, что он упал, ушибся и

нуждается в помощи. Для других участников ситуация более неопределенная: экспериментатор после очевидного падения ничего не говорит. При каждом условии исследуется по 20 участников. Экспериментатор фиксирует, сколько времени прошло (в секундах), прежде чем испытуемые предложили свою помощь.

0 наблюдателей, неопределенность 24;
2 наблюдателя, неопределенность 38;
0 наблюдателей, определенность 14;
2 наблюдателя, определенность 14.

2. Исследователь хочет проверить гипотезу о том, что на изучение лабиринта могут воздействовать размер подкрепления и его задержка. Животные по 40 раз проходят лабиринт; при каждой попытке записывается количество ошибок. Изучается шесть условий. В конце лабиринта находится:
   1) 10 мг еды, которые сразу выдают животному, или
   2) 10 мг еды, но, дойдя до цели, животное должно ждать 10 с, прежде чем получит еду, или
   3) 10 мг еды, но, дойдя до цели, животное должно ждать 20 с, прежде чем получит еду, или
   4) 20 мг еды, которые сразу выдают животному, или
   5) 20 мг еды, но, дойдя до цели, животное должно ждать 10 с, прежде чем получит еду, или
   6) 20 мг еды, но, дойдя до цели, животное должно ждать 20 с, прежде чем получит еду.

Среднее количество ошибок у животных в шести группах следующее:

1:10 2:40 3:45 4:5 5:35 6:40

Упражнение 8.3. Оценка количества участников

Для каждого из следующих планов с помощью имеющейся информации определите количество участников, которые могут потребоваться для проведения исследования (подсказка: для одного из планов это нельзя определить, не получив дополнительной информации).

1. Смешанный факторный план 3x3; для заполнения ячейки A1B1 требуется 10 участников.
2. Факторный план 2 x 3 с повторяемыми измерениями; для заполнения ячейки A1B1 требуется 20 участников.
3. Факторный план 2 x 2 x 2 с независимыми группами; для заполнения ячейки A1B1C1 требуется 5 участников.
4. Смешанный факторный план 2x4; для заполнения ячейки A1B1 требуется 8 участников.

ГЛАВА 9

Корреляционные исследования

Обзор задач главы

Вы только что завершили изучение четырех глав, посвященных экспериментальному методу в психологии. В оставшихся четырех главах рассматриваются другие методы исследований. Данная глава сосредоточена на корреляционном исследовании, связанном с изучением взаимосвязей между переменными. Далее вы узнаете, что интерпретацию корреляционных исследований необходимо проводить очень осторожно, а также, что на основании таких исследований можно предсказывать будущее поведение. Изучив эту главу, вы:

- поймете, что корреляционные исследования берут свое начало в работах Фрэнсиса Гальтона, и осознаете значимость речи Ли Кронбаха о «двух дисциплинах»;
- научитесь отличать положительные и отрицательные двумерные корреляции и отображать их с помощью графика рассеяния, а также узнаете, какие факторы могут влиять на значение коэффициентов корреляции (например, ограничение диапазона);
- научитесь вычислять коэффициент детерминации и интерпретировать его значение;
- поймете, как с помощью регрессионного анализа предсказывать будущее поведение;
- поймете, каким образом направленность может затруднить интерпретацию корреляции и как эта проблема решается с помощью корреляционной матрицы;
- научитесь понимать проблему третьей переменной и научитесь оценивать такую переменную с помощью процедуры взаимной корреляции;
- сможете описать различные ситуации исследований, в которых использование корреляционного подхода наиболее вероятно;
- сможете описать алгоритмы многомерных процедур множественной регрессии и факторного анализа.

Помните Роберта Вудвортса и его колумбийскую библию — первый в своем роде учебник по экспериментальной психологии (см. главу 5)? В своей книге он

впервые показал ставшие для нас привычными различия между независимыми и зависимыми переменными в экспериментальных исследованиях. Он также разделил методы на экспериментальные и корреляционные, что оказало сильное влияние на психологические исследования. Согласно Вудвортсу, в экспериментальном методе управляют переменными, а в корреляционном — «измеряется одна или несколько характеристик одного и того же человека [и] вычисляется корреляция между этими характеристиками...» (Woodworth, 1938, p. 3). Вудвортс постарался убедить читателей, что оба подхода одинаково важны. По его словам, корреляционный метод «необходимо отличать от экспериментального, [но] он скорее равен последнему по значению, нежели стоит выше или ниже...» (Woodworth, 1938, p. 3). Однако, сделав такое утверждение, Вудвортс предоставил читателям искать информацию о корреляционных исследованиях в других источниках и посвятил оставшиеся 820 страниц своей книги исследованиям, иллюстрирующим экспериментальный метод. Можно понять читателей, если они решат, что корреляционные исследования не так важны, как экспериментальные.

Две психологические дисциплины

В книге Вудвортса было положено начало разделению, которое в конечном счете привело к президентскому обращению Ли Кронбаха к Американской психологической ассоциации, озаглавленному «Две дисциплины научной психологии» *(Two Disciplines of Scientific Psychology)* (Cronbach, 1957). Как вы можете догадаться, две дисциплины — это корреляционная и экспериментальная психология. Согласно Кронбаху, корреляционная психология посвящена исследованию взаимосвязей между естественными переменными и изучению индивидуальных различий. Психологи-экспериментаторы, в свою очередь, обычно не изучают индивидуальные различия, а сводят их к минимуму или контролируют их с целью показать, что некоторые факторы предсказуемым образом влияют на поведение всех индивидуумов. Ученые, посвятившие себя корреляционным исследованиям, наблюдают переменные и выявляют взаимосвязи между ними, а экспериментаторы управляют переменными и наблюдают результаты. Первые ищут отличия людей друг от друга, а вторые стремятся открыть общие законы, распространяющиеся на всех людей.

Кронбах опасался, что корреляционное направление будет недооценено в научной психологии, и выразил надежду на синтез двух направлений. По его мнению, это должно привести к тому, что эти подходы будут одинаково высоко оценены сторонниками каждого из направлений, а исследования будут проводиться с учетом обеих стратегий. По его словам,

> далеко не достаточно, чтобы каждая из дисциплин заимствовала отдельные подходы у другой. Корреляционная психология изучает только различия между организмами: экспериментальная психология изучает только различия между воздействиями. Общая дисциплина будет изучать и то и другое, но кроме того, она будет учитывать, в противном случае забытые, *взаимодействия между переменными, связанными с организмами, и переменными воздействия.*
>
> Cronbach, 1957, p. 681, курсив мой. — *Дж. Г.*

Как свидетельствуют слова, выделенные курсивом, Кронбах требовал увеличения числа экспериментальных планов P x E (P — индивидуум или организм; E — среда или внешнее воздействие), с которыми вы познакомились в предыдущей главе. Также он требовал переоценки корреляционного метода в целом, что и произошло в последние 40 лет. Благодаря скорости и возможностям современных компьютеров такие сложные корреляционные процедуры, как множественная регрессия и факторный анализ сегодня широко распространены. Однако авторы многих учебников по экспериментальной психологии придерживаются созданной Вудвортсом традиции и уделяют мало внимания корреляционному методу, зачастую выделяя ему лишь одну или две страницы в главе о «неэкспериментальных» методах. Как вы увидите, эта книга является исключением. Но прежде, чем погружаться в суть корреляционных исследований, прочитайте вставку 9.1, в которой описано, как в ходе исследования сэра Фрэнсиса Гальтона, посвященного вопросам наследования гениальности, родилась корреляционная процедура.

ВСТАВКА 9.1

История — исследование гениальности Гальтона

Первое знакомство с сэром Фрэнсисом Гальюном (1822-1911), состоявшееся в главе 1, показало его как человека несколько эксцентричного (старающегося измерить пользу молитвы). Но вы совершите ошибку, если будете думать о Гальтоне, как о чудаке. Он был пионером эмпирических исследований интеллекта и одним из первых утверждал, что гениальность наследуется, а не возникает в результате воспитания. Кроме того, он открыл понятие корреляции.
На Гальтона произвела большое впечатления теория эволюции Дарвина, а в особенности мысль о том, что особи, принадлежащие к одному биологическому виду, отличаются друг от друга. Индивидуальные особенности, способствующие выживанию, подвергаются «естественному отбору» и передаются потомкам. Гальтон считал, что интеллект является особенностью, которая различается у всех людей, важна для выживания и наследуется так же, как физические характеристики, например как цвет глаз или рост. Он собрал факты, подтверждающие наследуемость интеллекта, и опубликовал две книги, посвященные этому вопросу: «Потомственные гении» (1869) и «Ученые-англичане: природа и воспитание» (1874). Последний труд популяризировал широко известные сегодня термины «природа» *(nature)* и «воспитание» *(nurture).* В своих работах Гальтон отметил статистическую тенденцию, заключающуюся в том, что гениальность и способности, проявляющиеся в определенных областях (например, способности к химии или юриспруденции), прослеживаются в нескольких поколениях внутри семьи. Однако он недооценил влияние окружающей среды и сделал вывод, что гениальность возникает в результате передачи наследственной информации. Он аргументировал свою точку зрения, в частности, тем, что интеллект в популяции имеет нормальное распределение. Другие наследуемые особенности (например, рост) также имеют нормальное распределение, и поэтому Гальтон принял этот статистический факт за показатель влияния наследственности (Fancher, 1990).

Только в 1888 г. ученому удалось показать высокую частоту появления таких черт, как гениальность в семьях: свои представления он сформулировал в работе, названной «Кореляция и ее измерение» (цитируется в Fancher, 1990). Во-первых, Гальтон обнаружил, что данные можно организовать по рядам и столбцам, как показано на рис. 9.1. Числа в каждой ячейке отражают количество людей, попадающих в категорию, определяемую заголовками рядов и столбцов. Так, наибольшее число в этих ячейках показывает, что выборка содержит 14 детей ростом от 67 до 68 дюймов, чьи родители также имели рост от 67 до 68 дюймов. Как вы вскоре узнаете, таблица Гальтона - это прототип сегодняшнего «графика рассеяния».

Рост детей (среднее арифметическое = 68 дюймов)

Рост родителей (среднее арифметическое = 68,1 дюйма)	63"–64"	64"–65"	65"–66"	66"–67"	67"–68"	68"–69"	69"–70"	70"–71"	71"–72"	72"–73"
72"–71"						1	2	2	2	1
71"–70"				2	4	5	5	4	3	1
70"–69"	1	2	3	5	8	9	9	8	5	3
69"–68"	2	3	6	10	12	12	2	10	6	3
68"–67"	3	7	11	13	14	13	10	7	3	1
67"–66"	3	6	8	11	11	8	6	3	1	
66"–65"	2	3	4	6	4	3	2			
Средний рост родителей в каждом столбце	67,2	67,3	67,4	67,6	67,9	68,2	68,4	68,8	69,1	69,3

Рис. 9.1. Таблица, построенная Гальтоном, чтобы показать корреляцию роста родителей и детей - прототип графика рассеяния

Во-вторых, Гальтон заметил, что когда «корреляция» была неполной, начинала проявляться одна закономерность. У родителей с ростом выше среднего были высокие дети, но довольно часто они были не такими высокими, как мать и отец. У родителей с ростом ниже среднего дети были низкие, но не настолько. Это означает, что рост у детей имеет тенденцию смещаться, или регрессировать, в сторону среднего арифметического значения в популяции. Феномен «регрессии к среднему», который, как вы уже знаете, представляет угрозу внутренней валидности исследования, является одним из самых выдающихся открытий Гальтона.

Третье наблюдение Гальтона состояло в том, что график, построенный по значениям среднего арифметического для каждого столбца таблицы рассеяния, дает более или менее прямую линию. По сути, он представляет собой разновидность «линии регрессии», с которой вы скоро познакомитесь. Таким образом, Гальтон открыл основные характеристики корреляционного анализа.

Прочитав о работе Гальтона, Карл Пирсон - молодой британский ученый, занимающийся статистикой, - продолжил изыскания в этой области и разработал формулу для вычисления коэффициента корреляции. Он обозначил коэффициент буквой «r», что значит «регрессия», в честь сделанного Гальтоном открытия регрессии к среднему. Вслед за Гальтоном Пирсон считал, что корреляционный анализ подтверждает идею о наследуемости многих свойств, обнаруживающих себя в отдельных семьях. Как вы вскоре узнаете, делать выводы о наличии причинно-следственных связей на основании корреляционного анализа (как поступили Гальтон и Пирсон) - весьма рискованное занятие.

Корреляция и регрессия: основы

Считается, что переменные коррелируют, если между ними существует какая-либо взаимосвязь. Это подразумевает сам термин «корреляция»: «ко» означает взаимное действие, а «реляция» (от англ. *relation)* — связь, отношение. В случае прямой или **положительной корреляции** взаимосвязь такова, что высокие значения одной переменной связаны с высокими значения другой, а низкие значения первой —

с низкими значениями второй. **Отрицательная корреляция** означает обратную взаимосвязь. Высокие значения одной переменной связаны с низкими значениями другой, и наоборот.

Положительная и отрицательная корреляция

Взаимосвязь между временем, посвященным занятиям, и оценками является примером положительной корреляции. Приведенные ниже данные, полученные в ходе гипотетического исследования восьми студентов, говорят о наличии положительной корреляции. В данном случае первой переменной является время, операционально определенное как количество часов в неделю, потраченных на учебу, а второй — средний балл (СБ), варьирующийся от 0,0 до 4,0.

	Время, потраченное на учебу	СБ
Студент 1	42	3,3
Студент 2	23	2,9
Студент 3	31	3,2
Студент 4	35	3,2
Студент 5	16	1,9
Студент 6	26	2,4
Студент 7	39	3,7
Студент 8	19	2,5

Значительное время, потраченное на учебу (42 часа), связано с высоким средним баллом (3,3), а самое малое время (16 часов) — с низким баллом (1,9).

Примером отрицательной корреляции может быть взаимосвязь между бесполезно потраченным временем и средним баллом. Бесполезно потраченное время можно операционально определить как количество часов в неделю, потраченное на определенные занятия, например на игру в видеоигры, просмотр телесериалов или игру в гольф (конечно, эти виды деятельности можно назвать и «терапией»). Ниже приведены гипотетические данные для других восьми студентов. На этот раз вы увидите обратную взаимосвязь между количеством часов в неделю, потраченных впустую, и средним баллом:

	Время, потраченное па учебу	СБ
Студент 1	42	1,8
Студент 2	23	3,0
Студент 3	31	2,2
Студент 4	35	2,9
Студент 5	16	3,7
Студент 6	26	3,0
Студент 7	39	2,4
Студент 8	19	3,4

Обратите внимание, что при отрицательной корреляции переменные имеют обратную взаимосвязь: большое количество потраченного зря времени (42) связано с низким средним баллом (1,8), а небольшое (16) — с более высоким (3,7).

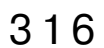
Силу корреляции показывает особая величина описательной статистики, носящая название «коэффициент корреляции». Коэффициент корреляции равен -1,00 в случае прямой отрицательной корреляции, 0,00 при отсутствии взаимосвязи и +1,00 при полной положительной корреляции. Наиболее распространенным коэффициентом корреляции является **пирсоново r** (о нем упоминалось во вставке 9.1), названное так в честь британского ученого, соперничающего в известности с сэром Рональдом Фишером. Пирсоново r вычисляется для данных, полученных с помощью интервальной шкалы или шкалы отношений. В случае других шкал измерений рассматриваются другие виды корреляции. К примеру, для порядковых данных (т. е. упорядоченных) вычисляется «ро» Спирмена. В приложении *С* показано, как вычислять пирсоново r.

Так же как среднее арифметическое и стандартное отклонение, коэффициент корреляции является величиной описательной статистики. В ходе заключительного анализа определяется, является ли конкретная корреляция значимо большей (или меньшей) нуля. Таким образом, для корреляционных исследований нулевая гипотеза (H_0) говорит, что действительное значение r равно 0 (т. е. нет никаких взаимосвязей), а альтернативная гипотеза *(H_1)* — что r № 0. Отвергнуть нулевую гипотезу — значит решить, что между двумя переменными существует значимая взаимосвязь. В приложении *С* показано, как определить, является ли корреляция статистически значимой.

График рассеяния

Силу корреляции можно обнаружить, рассмотрев современную версию построенной Гальтоном таблицы (рис. **9.1) — график рассеяния.** Он является графическим отображением взаимосвязи, на которую указывает корреляция. Как показано на рис. 9.2, в случае полной положительной (9.2, а) или полной отрицательной (9.2, б) корреляции точки образуют прямую линию, а нулевая корреляция дает график рассеяния (9.2, в), точки которого распределены случайным образом. По сравнению с относительно слабой корреляцией (9.2, г и д) точки сравнительно сильной расположены ближе друг к другу (9.2, ж и з). В целом, по мере ослабления корреляции точки на графике рассеяния все больше удаляются от диагонали, связывающей точки при полной корреляции, равной +1,00 или -1,00.

На рис. 9.3 показано, как по набору данных создается график рассеяния, а на рис. 9.4 приведены графики рассеяния для гипотетических примеров со средними баллами. Они отображают сильную положительную корреляцию между временем, потраченным на учебу, и средним баллом, а также сильную отрицательную корреляцию между бесполезно потраченным временем и средним баллом. Значения коэффициента корреляции равны +0,88 и -0,89 соответственно. Преподаватель может попросить вас проверить эти значения пирсонова r с помощью процедуры, приведенной в приложении *С.*



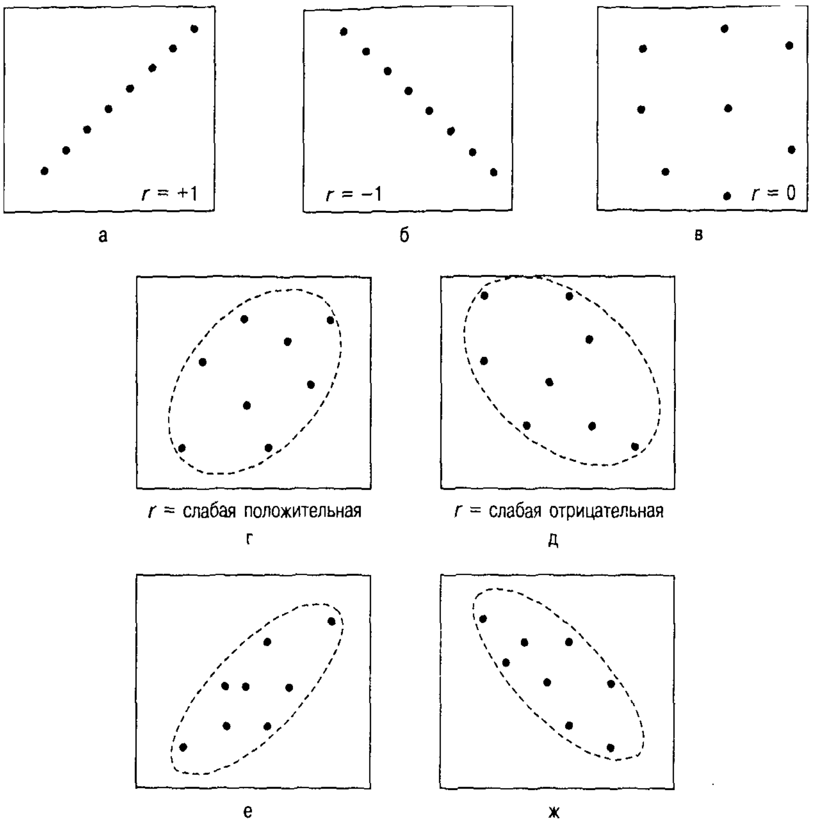


Рис. 9.2. Разновидности графиков рассеяния

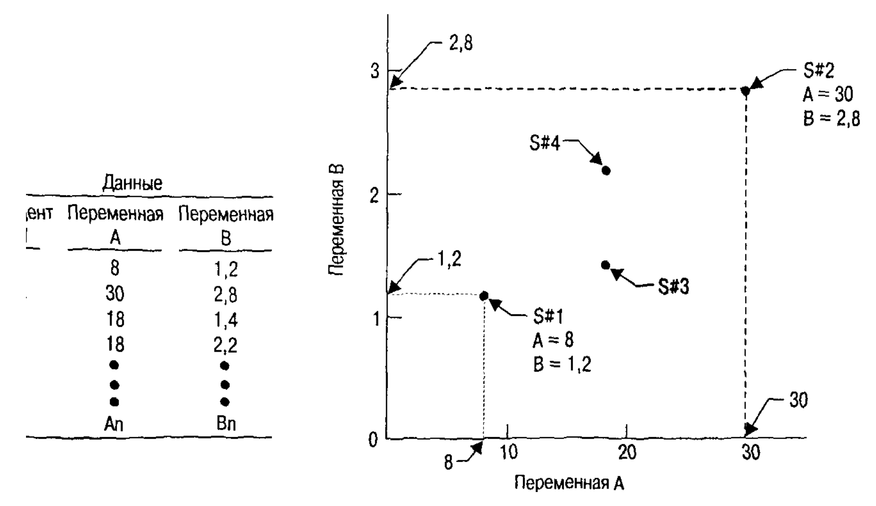


Рис. 9.3. Построение графика рассеяния для набора данных

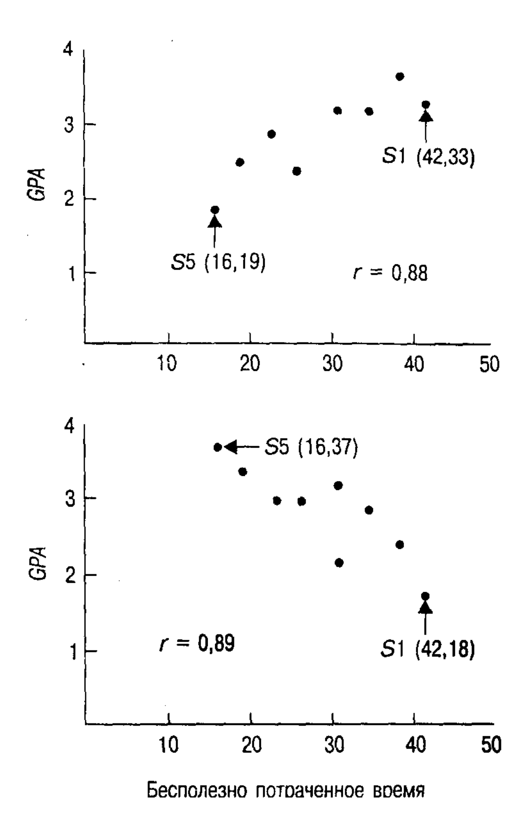


Рис. 9.4. Графики рассеяния для гипотетических данных о средних баллах

Допущение линейности

До сих пор изучаемые нами графики рассеяния состояли из точек, несколько отклоняющихся от прямой линии, которая образуется при полной корреляции со значениями -1,00 или +1,00. Однако не все взаимосвязи линейны, а вычисление пирсонова r для нелинейного случая не поможет выявить природу такой взаимосвязи. На рис. 9.5 показан гипотетический пример, отражающий одно известное психологическое открытие: взаимосвязь между возбуждением и выполнением задания. Сложные задания выполняются хорошо при среднем уровне возбуждения, но гораздо хуже при очень низком или очень высоком (см., например, Anderson, 1990). При очень низком уровне возбуждения у человека недостаточно сил, чтобы работать над заданием, а очень высокое возбуждение мешает эффективной обработке информации, требующейся для выполнения работы. Из графика рассеяния видно, что точки ложатся вдоль определенной кривой, но при попытке применить линейную корреляционную процедуру вы получите, что r равно нулю или очень близко к нему. Анализ криволинейных взаимосвязей, подобных изображенной на рис. 9.5, проводится особыми методами, рассмотрение которых не входит в задачу данной книги.

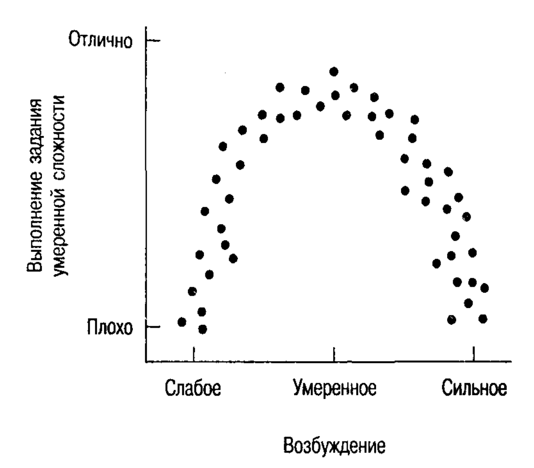


Рис. 9.5. Криволинейная взаимосвязь между уровнем возбуждения **и выполнением задания**

Ограничение диапазона

При проведении корреляционного исследования важно учитывать людей, оценки которых попадают в широкий диапазон. **Ограничение диапазона** одной или обеих переменных снижает корреляцию. Подобный эффект показан на рис. 9.6. Предположим, вы изучаете взаимосвязь между оценками *SAT* (американский школьный тест проверки способностей) и успеваемостью в колледже (последняя оценивается по средним баллам, полученным первокурсниками в конце года). На рис. 9.6, а показано, каким может быть график рассеяния при исследовании 25 студентов. Коэффициент корреляции равен +0,70. Допустим далее, что вы решили изучить эту взаимосвязь на примере студентов, получивших 1200 и более баллов по тесту *SAT* На рис. 9.6, б выделены точки графика рассеяния для таких студентов — как показано на рис. 9.6, в, по ним можно построить отдельный график. Если вы теперь сравните рис. 9.6, а и 9.6, в, вам станет ясно, что для 9.6, в корреляция ниже. Действительно, она падает до +0,26.

Этот пример имеет интересное следствие для колледжей, которые не учитывают студентов, чьи суммарные оценки теста *SAT* меньше 1200 баллов. Различные исследования (например, Schrader, 1971) показали, что общая корреляция между оценками *SAT и* баллами, полученными в конце первого курса, равная приблизительно +4,0, является статистически значимой, но не высокой. Корреляцию находили, используя студентов с самыми разными оценками теста *SAT.* Если диапазон оценок *SAT* ограничен 1200 баллами и выше, то корреляция заметно снижается. Существуют особые процедуры для «коррекции» корреляции с учетом проблемы ограничения, но необходимо осознавать, что ограничение диапазона непосредственно влияет на возможность строить предположения о дальнейших событиях. Учебные заведения, проводящие строгий отбор и отсеивающие абитуриентов с результатами теста *SAT* ниже 1200 баллов, без сомнения получат хороших студентов, но их возможность предсказать учебную успеваемость на основании оценок *SAT* будет не такой высокой, как у заведений, не предъявляющих таких высоких требо-

ваний к абитуриентам. В вузах, имеющих меньше ограничений по отбору студентов, корреляция оценок *SAT* и академической успеваемости будет выше, чем в учебных заведениях со строгими ограничениями.

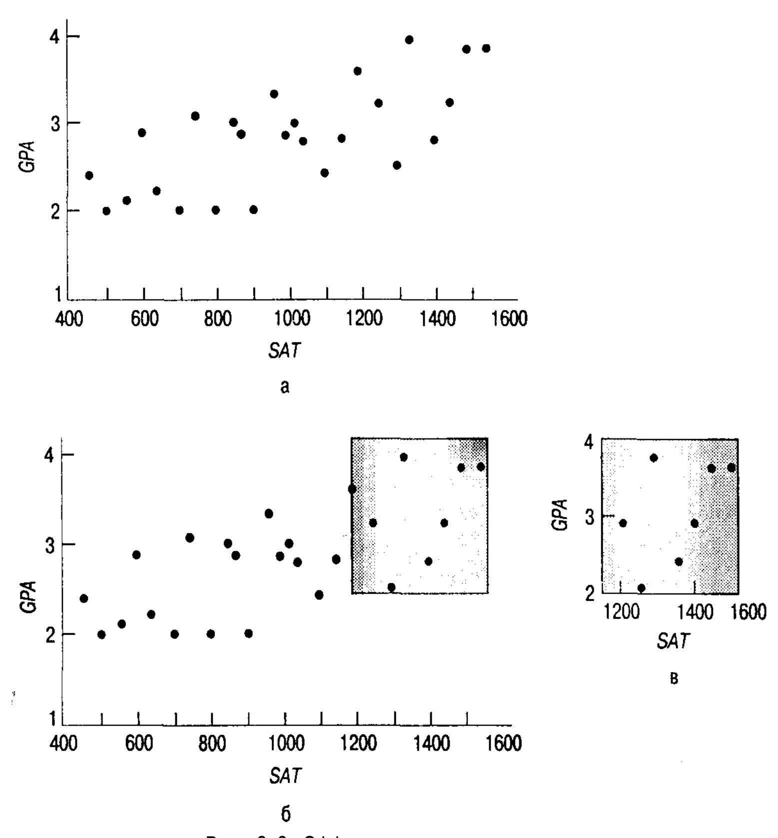


Рис. 9.6. Эффект ограничения диапазона

Коэффициент детерминации — r^2

Довольно легко неверно понять смысл конкретного значения пирсонова r. Если оно равняется +0,70, то взаимосвязь действительно является относительно сильной, но студенты иногда думают, что +0,70 каким-то образом связано с 70%, и считают, что в таком случае взаимосвязь установлена на 70%. Это неверно. Для интерпретации значения корреляции гораздо правильнее использовать **коэффициент детерминации (r^2).** Он находится возведением в квадрат пирсонова r, а поэтому вне зависимости от типа корреляции (положительной или отрицательной) его значение никогда не бывает отрицательным. Данный коэффициент формально определяется как степень изменчивости одной переменной корреляции, вызванная изменчивостью другой переменной. Поясним это на конкретном примере.

Предположим, вы проводите исследование, в ходе которого у 100 участников измеряется уровень эмоциональной депрессии и средний балл. Вы проверяете вза-

имосвязь между двумя переменными и обнаруживаете отрицательную корреляцию: чем выше уровень депрессии, тем ниже средний балл, и наоборот, чем слабее депрессия тем выше средний балл. Рассмотрим два значения корреляции, которые могут быть получены в результате этого исследования, - -1,00 и -0,50. Коэффициент детерминации будет равен 1,00 и 0,25 соответственно. Чтобы понять смысл этих значений, для начала обратим внимание на то, что средний балл у 100 изучаемых людей, скорее всего, будет варьироваться от 0,0 до 4,0. Как исследователи, мы хотим выяснить причину такой изменчивости — почему один человек получает 3,8 балла, а другой 2,4 и т. д. Другими словами, мы хотим узнать, что вызывает индивидуальные различия в средних баллах? В действительности, причиной этому может быть несколько факторов: учебные привычки, общий уровень интеллекта, эмоциональная устойчивость, склонность к выбору легких предметов для изучения и т. д. Как показывают оценки теста на депрессию, в нашем гипотетическом исследовании изучается один из этих факторов — эмоциональная устойчивость, r^2 показывает, насколько изменчивость средних баллов может быть связана непосредственно с депрессией. В первом случае, когда $r = -1{,}00$, а $r^2 = 1{,}00$, мы можем прийти к выводу, что 100% изменчивости средних баллов связана с изменчивостью оценок депрессии. Следовательно, можно сказать, что 100% различий между средними баллами (3,8 и 2,4 и др.) вызваны депрессией. В реальном исследовании такой результат, конечно, невозможно получить. Во втором случае, когда $r = -0{,}5$, а $r^2 = 0{,}25$, только одна четверть (25%) изменчивости средних баллов будет связана с депрессией. Остальные 75% связаны с другими факторами, подобными перечисленным выше. Говоря кратко, коэффициент детерминации лучше характеризует силу отношений, чем пирсоново r.

Регрессионный анализ: построение предположений

Как показано в приведенном выше примере, важнейшей особенностью корреляционных исследований является возможность при наличии сильной корреляции строить предположения о будущем поведении. Корреляция между двумя переменными дает возможность на основании значений одной из них предсказать значения другой. Это несложно показать на примере со средними баллами. Если вы знаете, что время, посвященное учебе, и средний балл коррелируют, и я скажу вам, что некто занимается 45 часов в неделю, вы сможете безошибочно предсказать относительно высокий средний балл для такого студента. Аналогично высокий средний балл позволит вам предсказать время, уделяемое учебе. Как вы узнаете позднее в этой главе, корреляционные исследования обеспечивают базу для построения предположений по результатам психологических тестов. Построение предположений на основании корреляционных исследований называется **регрессионным анализом.**

На рис. 9.7 представлен график рассеяния для: а) времени, посвященного учебе и среднего балла и б) бесполезно потраченного времени и среднего балла, но на этот раз на каждом графике отображена **линия регрессии,** которая используется для построения предположений. Линию регрессии также называют «оптимальной линией»: она представляет собой наилучший из возможных способов обобщения точек графика рассеяния. Это значит, что абсолютные значения расстояний по вертикали между каждой точкой графика и линией регрессии минимальны.

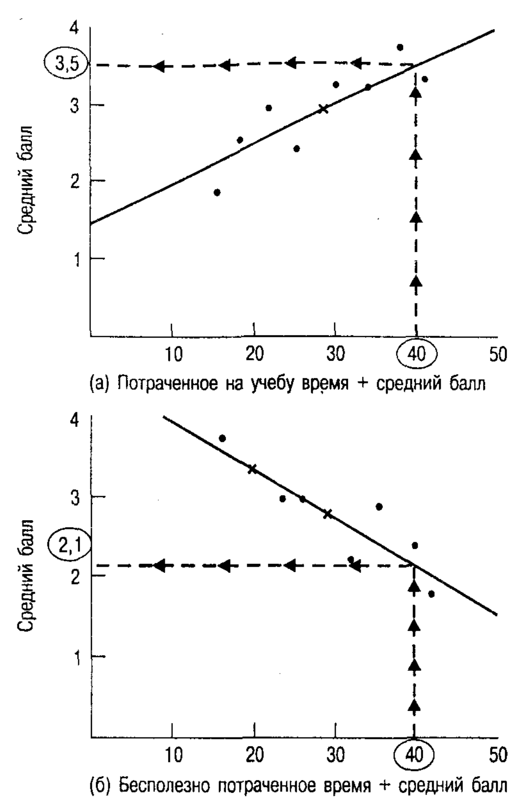


Рис. 9.7. Графики рассеяния с линиями регрессии

Линия регрессии рассчитывается по той же формуле, по **которой в школе вы** строили прямую по заданным координатам *X* и *Y*:

$$Y = a + bX,$$

где *a* — это точка, в которой прямая пересекает ось Y(т. е. отрезок, отсекаемый на оси Y), а *b* — это угол наклона прямой, или ее относительная крутизна. *X* — это известная величина, а Y— величина, которую вы пытаетесь предсказать. Зная 1) силу корреляции и 2) стандартное отклонение для коррелирующих переменных, можно вычислить величину b, а зная 1) значение b и 2) средние значения коррелирующих переменных, можно найти a. В приложении *С* на конкретном примере показано, как вычислять эти значения и подставлять их в уравнение регрессии.

В регрессионном анализе для предсказания значения Y(например, среднего балла) на основании значения *X* (например, времени, посвященного учебе) используется уравнение регрессии. У иногда называют **критериальной переменной,** а *X* — **предикторной переменной.** Однако для построения точных предположений корреляция должна быть значительно выше нуля. Чем выше корреляция, тем ближе будут точки графика рассеяния к линии регрессии и тем больше будет уверенность в том, что ваши предположения верны. Таким образом, отмеченная ранее проблема ограничения диапазона, которая снижает корреляцию, также снижает достоверность предсказаний.

Из рис. 9.7 также видно, как строить предположения с помощью линии регрессии. Поскольку между временем, отведенным на учебу, и средним баллом существует взаимосвязь, можно спросить, какой средний балл стоит ожидать у студента, который проводит за учебой по 40 часов в неделю. Чтобы представить процесс наглядно, проведем пунктирную линию от оси *X* к линии регрессии, а затем повернем влево на 90° и проведем линию до пересечения с осью *Y*. Значение точки на оси У и будет предполагаемым значением (помните, что правильность предположения зависит от силы корреляции). Таким образом, по времени учебы, равному 40 часам, можно предсказать средний балл, равный 3,5, а по бесполезно потраченным 40 часам — средний балл чуть выше 2,1. С помощью формулы регрессии можно вычислить более точные значения: 3,48 и 2,13 соответственно, и сделать более точные предсказания. Еще раз см. приложение *С*.

Будьте уверены, что регрессионный анализ применяется в большинстве исследований, о которых вы узнаете из средств массовой информации. К примеру, вам может встретиться отчет об исследовании «факторов риска для инфаркта», в котором на основании значимой корреляции, скажем между курением и сердечными заболеваниями, сделан вывод, что у людей, злоупотребляющих курением, больше вероятность развития сердечно-сосудистых заболеваний, чем у некурящих. Это значит, что курение является основанием для предсказания развития болезней сердца. На основании другого исследования, посвященного изучению «портрета жестокого супруга (супруги)», может быть сделан вывод о том, что вероятность появления подобного поведения увеличивается, если виновник — безработный. Это следует из наличия корреляции между безработицей и склонностью к оскорбительному поведению. На основании наличия корреляции с помощью регрессионного анализа, зная первое, можно сделать предположение насчет второго.

Последнее замечание по поводу регрессионного анализа касается и практической, и этической сторон дела. Предположения могут распространяться только на тех людей, чьи оценки попадают в диапазон, на котором основана корреляция. Например, если уравнение регрессии, предсказывающее хорошую успеваемость в колледже, основано на исследовании, проведенном с участием белых испытуемых, живущих в провинции, принадлежащих к среднему классу и имеющих оценки, варьирующиеся от 1000 до 1400, то на основании этого уравнения нельзя предсказать успеваемость других абитуриентов, не принадлежащих к данной популяции.

Интерпретация корреляций

Корреляционные исследования являются мощным инструментом, дающим возможность предсказывать будущие события. Однако необходимо очень внимательно подходить к интерпретации результатов таких исследований. Обнаружение корреляции между двумя переменными не позволяет сделать вывод, что одна переменная является причиной появления другой. К сожалению, недостаток понимания этого общего правила приводит к тому, что в корреляционных исследованиях люди разбираются хуже всего. В новостях нередко описываются корреляционные исследования, и некритичным читателям сообщают, что одна из взаимосвязанных переменных является причиной другой. Например, было проведено огромное ко-

личество исследований, показавших, что определенный вид облысения является признаком будущего сердечного заболевания. В одном из них (Lotufo, Chae, Ajani, Hennekens, & Manson, 1999) исследователи опросили более 22 000 врачей и обнаружили, что у людей с коронарной болезнью часто наблюдается облысение макушки головы (гораздо чаше, чем облысение передней части головы или общее выпадение волос). В большинстве отчетов в средствах массовой информации было указано на отсутствие причинно-следственной связи, но некритичные читатели вполне могли решить, что облысение — это верный признак будущего инфаркта. Я не знаю, увеличилось ли количество операций по трансплантации волос после того, как был опубликован отчет об этом исследовании, по все те, кто думают, что, потратив деньги на трасплаптант, они избавятся от риска получить инфаркт, являются жертвами ложной интерпретации корреляционного исследования.

Если в подобном образе мысли вы усматриваете проблему, то у вас уже есть чувство, которое я стараюсь в вас развить. Как вы помните из главы 1, одна из моих задач состоит в том, чтобы помочь вам стать более критичными потребителями информации. В публикуемых описаниях медицинских и психологических корреляционных исследований довольно часто встречаются ложные выводы и знание того, как правильно интерпретировать корреляционные исследования, поможет вам адекватно отнестись к таким работам.

Корреляция и причинно-следственная связь

Как вы уже знаете, в экспериментальном исследовании с управляемой независимой переменной с достаточной степенью уверенности можно делать выводы о наличии причинно-следственных связей. Изучаемая переменная является управляемой, и если все остальные факторы поддерживаются постоянными, то результат можно отнести непосредственно на счет независимой переменной. Однако в случае корреляционного исследования невозможно поддерживать постоянными другие факторы, а следовательно, нельзя делать выводы о причинах и следствиях. Рассмотрим две проблемы интерпретации, возникающие в корреляционных исследованиях. Это проблема обратимости, или направленности, и проблема третьей переменной.

Направленность

Если между двумя переменными *A* и *B* существует корреляция, то может случиться так, что *A* вызывает *B (A > B),* а *B* вызывает А *(B > A).* Действие причинно-следственной связи в двух направлениях называется **проблемой направленности.** Существование корреляции *само по себе* еще ничего не говорит о направлении этой связи.

Исследования взаимосвязи между просмотром телевизора и детской агрессией дают типичный пример проблемы направленности. Некоторые из этих исследований являются корреляционными и проводятся следующим образом. Измеряется время просмотра телевизора (переменная *A),* например количество часов в неделю, уделяемое этому занятию. Далее у тех же детей измеряется уровень агрессии (переменная *B).* Он может представлять собой обобщенную оценку агрессивности, данную учителями этим детям. Предположим, в таком исследовании была обна-

ружена корреляция +0,58, что значимо превышает нулевое значение. Какой вывод можно сделать на основании таких данных?

Можно, конечно, заключить, что поскольку длительный просмотр телевизора неизбежно приводит ребенка к восприятию большого количества сцен с насилием, и, как мы знаем, дети учатся в процессе наблюдения, то длительный просмотр телевизионных передач приведет к тому, что дети станут агрессивными. Это означает $A > B$.

Но будет ли причинно-следственная связь проявляться в обратном направлении? Возможно ли, что агрессивные дети просто любят смотреть телевизор больше, чем их неагрессивные товарищи? Возможно, зная о существовании большого количества передач, включающих сцены насилия, агрессивные дети захотят проводить больше времени за занятием, которое их по-настоящему интересует. Таким образом, возможно, что агрессивность является причиной просмотра большего количества передач со сценами насилия. Это означает $B > A$.

Если единственным основанием является наличие корреляции, то правильно определить направленность причинно-следственной связи невозможно. Однако есть способ решить проблему направленности. Он основан на критериях определения причинно-следственной связи, о которых было рассказано в главе 1. Как вы помните, психологи-исследователи считают, что причинно-следственная связь между *A* и *B* установлена, если эти события появляются вместе, если *A* предшествует *B* во времени, если эта связь вытекает из некоторой теории или если можно опровергнуть другие объяснения их совместного появления.

В случае описанного ранее исследования просмотра телевизионных передач и агрессии мы знаем только, что *АиВ* появляются вместе, а на основании информации о возможности обучения в процессе наблюдения разумно предположить наличие причинно-следственной связи между ними. Но с большей уверенностью определить направленность взаимосвязи можно с помощью процедуры, носящей название **перекрестной корреляции.** Эта процедура состоит в изучении корреляции между переменными в различные моменты времени, а следовательно, она представляет собой разновидность лонгитюдного экспериментального плана. Следующий широко известный пример иллюстрирует данную процедуру.

Пример 19. Телевидение и агрессия

Эрон с соавторами (Eron, Huesman, Lefkovitz, & Walder, 1972) занимались исследованием взаимосвязи между просмотром телевидения и агрессией, которой был посвящен наш гипотетический пример. В частности, они измеряли: а) склонность к просмотру телепередач с насилием и б) сопутствующую ей агрессивность. В 1960 г. ученые исследовали 875 третьеклассников, проживающих в сельской местности штата Нью-Йорк. Между просмотром телепередач с насилием и агрессивностью была обнаружена умеренная, но значимая корреляция +0,21, Но самое интересное, что через 10 лет Эрон и его соавторы нашли 427 человек из прежних участников (теперь ставших «тринадцатиклассниками») и оценили у них значения тех же двух переменных. По данным измерения двух переменных в два различных момента времени может быть вычислено шесть корреляций. Данные о корреляциях, полученные Эроном и др. в ходе исследования, показаны на рис. 9.8.

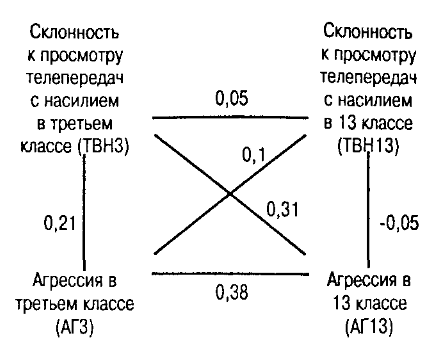


Рис. 9.8. Результаты перекрестного исследования влияния склонности к просмотру телепередач с насилием на последующую агрессию

Особый интерес представляют «перекрестные» корреляции, так как они отображают взаимосвязи между переменными, разделенными во времени. Если агрессивность третьеклассников вызывает позднейшую склонность к просмотру телепередач с насилием *(В > А),* то мы должны ожидать наличие довольно высокой корреляции между агрессивностью в момент 1 и склонностью к определенным телепередачам в момент 2. В действительности корреляция практически равна нулю (+0,01). Если же ранняя склонность к просмотру телепередач с насилием приводит к позднейшему агрессивному поведению *(А > В),* то будет значительной корреляция между склонностью в момент 1 и агрессивностью в момент 2. Как видно из приведенных данных, эта корреляция равна +0,31 — не слишком высокий, но значимый показатель. На основании этих данных, а также другой информации, полученной в ходе исследования, Эрон и его коллеги сделали вывод о том, что ранняя склонность к просмотру телепередач с насилием является причиной позднейшей агрессивности.

Однако необходимо очень осторожно подходить к интерпретации перекрестной корреляции (Rogosa, 1980). Например, если вы внимательно изучите корреляции на рис. 9.8, то заметите, что корреляцию со значением +0,31 можно частично отнести на счет корреляций со значениями +0,21 и +0,38. Это значит, что вместо прямой зависимости между склонностью в 3-м классе и агрессией в 13-м эта связь определяется наличием взаимосвязей между склонностью к телепередачам с насилием и агрессией в 3-м классе и между двумя оценками агрессии. Ребенок, имеющий высокие показатели по склонности к передачам с насилием в 3-м классе, может быть агрессивным в этом возрасте и оставаться таким же агрессивным (или даже более агрессивным) в 13-м классе. И наоборот, возможно, что агрессивность в 3-м классе вызывает а) склонность к просмотру телепередач с насилием в 3-м классе и б) позднейшую агрессивность. Таким образом, перекрестная корреляция помогает решить проблему направленности, но не вопрос интерпретации. В более общих словах, трудности с интерпретацией проявляются как проблема третьей переменной.

Третья переменная

Выпуск «Нью-Йорк Таймс Мэгэзин» от 4 июня 2000 г. содержал веселую статью под названием «Мирные лужайки» (Plotz, 2000). Ее автор поднял важный вопрос о том, почему одни страны постоянно воюют, а другие ведут себя относительно

мирно. Его ответом был гольф — страны, в которых значительная часть популяции играет в гольф, менее воинственны, чем страны, не играющие в гольф. Плоц описал свою идею следующим образом:

> Все мирные европейские нации любят гольф. Но Россия, воюющая с Чечней, не входит в их число. Турция и Греция, не играющие в гольф, долгое время воевали из-за Кипра, также не играющего в гольф. Бывшая Югославия раскололась на пять государств. И только мирная Словения любит помахать клюшкой. Играют ли *в* гольф в Индии или Пакистане? Конечно, нет. Алжирцы убивают друг друга, а их соседи-марокканцы забивают мячи в лузы (р. 32).

Хотя лозунг «Играйте в гольф, а не в войну» (р. 37) заслуживает всяческого одобрения, я думаю, что вы понимаете абсурдность заявления, что гольф является причиной мира. И лишь немногим более вероятно, что верно обратное — мир является причиной игры в гольф. Если на уровне нации и есть корреляция между миром и игрой в гольф, то поскольку Плоц не представил пирсоново r, ее существование представляет собой гиперболизированный пример **проблемы третьей переменной.** Так как в случае корреляционного исследования контроль за внешними переменными не осуществляется, причина найденной корреляции может заключаться именно в них. Это означает, что А не является причиной *В* и *В* не является причиной Л, а некоторая неизвестная третья переменная *С* вызывает как *А,* так и *В. С* — это некая неконтролируемая третья переменная (или переменные — часто бывает, что за корреляцией стоит более одной неконтролируемой переменной). Попробуйте придумать третью переменную, вызывающую эту сомнительную корреляцию мира и игры гольф. Возможно, что благополучные страны имеют больше оснований быть мирными, а люди в них могут иметь больше времени для досуга, в том числе и для гольфа.

Другие примеры проблемы третьей переменной не настолько абсурдны, как пример с мирными странами и игрой в гольф, а поэтому они могут легко подтолкнуть людей к установлению несуществующих причинно-следственных связей, что может иметь весьма печальные последствия. Например, из вставки 9.1 вы узнали о разделяемом Гальтоном и Пирсом убеждении, что положительная корреляция между родителями и детьми в отношении определенных особенностей доказывает, что наследственность является первопричиной интеллектуальных различий. Оба ученых не учли влияния окружающей среды (например, родители с высоким уровнем интеллекта могут создавать стимулирующее окружение для своих детей) как фактора, на котором основывается взаимосвязь способностей родителей и детей. Интересно, что Гальтон был основателем «евгенического» движения, сторонники которого, кроме прочего, считали, что Великобритания должна путем строгого подбора супружеских пар увеличить общий уровень интеллекта своих граждан. Гальтон весьма цинично заметил, что как возможно «путем тщательного отбора получить породу собак или лошадей, обладающую необычной скоростью бега... так, заключая разумные браки на протяжении нескольких поколений, можно создать вид высокоодаренных людей» (Galton, 1869, процитировано в Fancher, 1990, р. 228). Несложно представить обратную сторону евгенического движения — мысль о том, что людям, считающимся «неподходящими» или «дефективными», следует запретить иметь детей. Именно такие мысли возникли у немецкого руководства в 1930-х гг.

Взаимосвязь между просмотром телепередач с насилием и детской агрессивностью является более современным примером проблемы третьей переменной. Как мы узнали, возможно, что передачи с насилием увеличивают агрессию *(А > В),* но также возможна причинно-следственная связь, имеющая обратную направленность *(В > А):* агрессивные дети могут иметь склонность к просмотру телепередач с насилием. Третья возможность заключается в том, что и *А* и *В* возникают как результат действия третьей переменной, *С (С > А + В).* К примеру, родители у таких детей могут быть жестокими. Дети, копируя агрессивное поведение родителей, также будут жестокими, и кроме того, подобное поведение родителей приведет к тому, что дети станут много смотреть телевизор, стараясь «сидеть тихо» и избегать контакта с родителями, которые их часто наказывают. Другой третьей переменной может быть отсутствие беглости речи. Возможно, дети становятся агрессивными из-за того, что их доводы в спорах не действуют, и они также много смотрят телевизор, чтобы избежать разговоров с окружающими.

Иногда попытки определить третью переменную являются делом сугубо умозрительным, а в других случаях могут быть причины подозревать влияние конкретной третьей переменно!!. В последнем случае, если вообще существует возможность измерить третью переменную, ее влияние можно оцепить с помощью процедуры, названной **взаимной корреляцией.** Данная процедура призвана осуществлять статистический контроль за третьей переменной. В действительности она представляет собой попытку *post .factum* (после свершившегося факта) создать по крайней мере полуэквивалентные группы. Предположим, к примеру, что вам известно, что между скоростью чтения и его осознанностью существует высокая корреляция, допустим +0,55 (пример из Sprinthall, 2000). Кроме того, вы подозреваете, что причиной этой корреляции является третья переменная, а именно *IQ*. Чтобы провести процедуру взаимной корреляции, вам потребуется найти корреляцию между а) *IQ* и скоростью чтения и б) *IQ* и осознанностью чтения. Допустим, что корреляция равна +0,70 и +0,72 соответственно. Значения достаточно высоки, чтобы заподозрить, что *IQ* является третьей переменной. Для вычисления взаимной корреляции необходимо использовать все три полученных значения корреляции (точную процедуру см. у Sprinthall, 2000). В результате будет получена взаимная корреляция, показывающая взаимосвязь между скоростью и осознанностью чтения, при которой значение *IQ* «контролируется». В данном случае взаимная корреляция равна +0,10. Таким образом, если *IQ* статистически контролируется («разделяется на части»), корреляция между скоростью и осознанностью практически исчезает, а это означает, что *IQ* действительно является важной третьей переменной, заметно влияющей на получение начальной корреляции между скоростью и осознанностью, равной +0,55.

В исследовании Эрона и др. (Eron et al., 1972), чтобы понять, существуют ли важные третьи переменные, приводящие к относительно высокой корреляции (+0,31) между склонностью 3-классников к насилию и агрессивностью 13-классников, было вычислено несколько взаимных корреляций. В табл. 9.1 показаны результаты, полученные для 12 различных потенциальных третьих переменных (в таблице они названы «контролируемыми переменными»). Как вы видите, значения взаимной корреляции варьируются от +0,25 до +0,31. Это говорит о том, что

ни один из 12 факторов не сделал значимого вклада в начальную корреляцию +0,31. Таким образом, даже при учете других факторов корреляция между ранней склонностью к просмотру телепередач с насилием и позднейшей агрессивностью *остается* близкой +0,31. Анализ подкрепил выводы о «наличии влияния просмотра телевизионных передач с насилием в ранние формирующие годы на позднейшую агрессию» (Егоп et al., 1972, р. 263).

Таблица 9.1

Взаимная корреляция в исследовании Эрона с соавторами (1972), посвященном изучению взаимосвязи между склонностью к просмотру телепередач с насилием и агрессивным поведением

Контролируемая переменная	Частичная корреляция между ТВНас3иАг13*
Отсутствует (начальная корреляция)	0,31
Переменные для третьеклассников	
Оценка агрессии другими школьниками	0,25
Позиция, занимаемая отцом на работе	0,31
IQ ребенка	0,28
Агрессивность отца	0,30
Агрессивность матери	0,31
Применяемые наказания	0,31
Надежды родителей, связываемые с ребенком	0,30
Мобильность родителей	0,31
Количество часов, проводимых у телевизора	0,30 -
Переменные для тринадцатиклассников	
Позиция, занимаемая отцом на работе	0,28
Собственные ожидания испытуемого	0,28
Количество часов, проводимых за телевизором	0,30

* ТВНас3 -= склонность к просмотру телепередач с насилием в 3-м классе; Аг13 = уровень агрессивности в 13-м классе.

Источник: Eron et al., 1972, табл. 4.

Внимание: корреляционная статистика или корреляционные исследования?

Прочитав предыдущие главы, посвященные экспериментальным планам, и эту главу, рассказывающую о корреляционных исследованиях, вы можете решить, что если в описании исследования встречается пирсоново r, то в этом исследовании обязательно должен использоваться «корреляционный план». К сожалению, все не так просто. Необходимо различать, когда корреляция является лишь частью статистического анализа, а когда используется корреляционный план. Пирсоново r можно вычислить в любом исследовании, в котором требуется изучить взаимосвязь между двумя переменными. Хотя чаще всего это делается в корреляционных

исследованиях, при проведении экспериментов также вычисляют пирсоново r, чтобы найти корреляцию между независимой и зависимой переменными. По сути, выявление силы эффекта (глава 4) равнозначно оценке уровня взаимосвязи между независимыми и зависимыми переменными.

Так же как присутствие пирсонова r в результатах само по себе не означает использования корреляционного плана, наличие критерия Стьюдента или дисперсионного анализа не означает, что был использован экспериментальный план. В качестве примера рассмотрим исследование, в ходе которого сравнивается степень подчинения власти у интровертов и экстравертов. Один из способов проведения такого исследования — это провести тест на интроверсию, затем тест на подчинение и проверить корреляцию между ними. Второй способ — это провести тест на интроверсию, затем отобрать участников, получивших 25% самых высоких (т. е. интровертов) и 25% самых низких (т. е. экстравертов) оценок. Далее этих испытуемых можно поставить в ситуацию, требующую подчинения, и определенным образом измерить степень демонстрируемого ими подчинения. Оценки подчинения, полученные двумя группами, можно сравнить, найдя коэффициент Стьюдента или проведя дисперсионный анализ. Но обратите внимание, что полученные результаты будут говорить о причине и следствии не больше, чем значение пирсонова r для двух переменных. Таким образом, вне зависимости от способа организации данных в корреляционных исследованиях существуют проблемы интерпретации.

В заключение следует отметить, что в корреляционных исследованиях возникают абсолютно такие же проблемы интерпретации, как и в исследованиях с субъективными независимыми переменными, описанных в предыдущих главах. В последних двух главах было описано три вида таких планов:

Однофакторпые планы с неэквивалентными группами (как двухуровневые, так и многоуровневые);

факторные планы с неэквивалентными группами;

факторные планы P x E.

Использование корреляций

Принимая во внимание подводные камни интерпретации корреляций, а также тот факт, что желаемые выводы о причинах и следствиях могут быть сделаны только на основании настоящих экспериментов с управляемыми независимыми неременными, почему не ограничиться проведением одних экспериментальных исследований? Зачем вообще необходимо вдаваться в сложности корреляционных исследований?

Необходимость корреляционных исследований

Ответ касается как практической, так и этической стороны дела. С практической стороны очевидно, что некоторые важные исследования просто невозможно было бы провести в виде экспериментов. При исследовании таких важных вопросов, как половые особенности поведения, различия между возрастными группами или типами личности невозможно провести случайное распределение. Изучение корреляции между тяжестью клинической депрессии и выполнением заданий имеет

смысл даже при условии, что субъектов нельзя распределить по условиям «легкая депрессия» и «тяжелая депрессия». Ли Кроибах в своей речи о двух дисциплинах сказал, что с помощью корреляционных исследовании «можно изучать то, что [мы] не можем или еще не научились контролировать» (Cronbach, 1957, p. 672). Кроме того, некоторые исследования имеют своей целью построение предположений о будущем поведении. Например, обнаружение корреляции между определенным свойством личности и успешностью в работе позволит специалистам по кадрам, не зная причин существования подобной взаимосвязи, предсказать, кто будет успешно выполнять конкретную работу.

Во-вторых, с этической точки зрения ясно, что есть исследования, которые невозможно провести в виде экспериментов с управляемыми переменными. Французский физиолог XIX века Поль Брока открыл мозговой центр речи, позднее названный его именем, заметив взаимосвязь между определенными видами нарушений речи и характером поражения мозга, обнаруженного при посмертном вскрытии (Goodwin, 1999). Для *экспериментального* подтверждения того, что причиной нарушения явилось поражение мозга, потребовалось бы случайным образом распределить людей па группу «с пораженным мозгом», имеющую нарушения структуры мозга, и более сохранную «контрольную» группу. В действительности потребовалась бы еще третья группа, которую можно обозначить как группу «с мнимым поражением мозга». Члены этой группы должны были бы подвергнуться большинству хирургических операций из перенесенных группой с поражением мозга, кроме действительного разрушения мозга. Я думаю, вы можете представить себе трудности отбора людей для такого эксперимента. В этом заключается одна из причин, по которым в экспериментальных исследованиях, посвященных изучению взаимосвязей между мозгом и поведением, используются животные.

Разновидности корреляционных исследований

Исследования, проводящиеся с помощью корреляционных процедур, можно обнаружить во всех областях психологии. Особенно часто они используются в исследованиях: а) посвященных разработке психологических тестов, б) по психологии личности и анормальной психологии и в) направленных на изучение врожденных особенностей и воспитания, начало которым было положено Фрэнсисом Гальтоном. Во всех этих исследованиях особое значение придается характеру отличий людей друг от друга. Далее мы рассмотрим по одному примеру для каждого из трех видов.

Пример 20. Психологическое тестирование

Из главы 4 вы узнали, что для того, чтобы измерение было значимым, оно должно быть надежным и валидным. Надежные измерения можно воспроизвести и они относительно свободны от ошибок измерения. Измерение валидно, если действительно измеряется та особенность, которая интересует. Надежное и валидное измерение интеллекта дает одинаковые значения *IQ* при двух различных попытках измерения и действительно отображает интеллектуальные способности, а не что-либо другое. Как показывает описанная ниже серия исследований, оценка надежности и валидности в большой мере зависит от корреляции.

Для измерения интеллекта у детей используются два теста: Стэнфордская редакция шкалы Бине—Симона и шкала Векслера для измерения интеллекта детей

(*WISC*). Однако недавно появился тест, названный Оценочная батарея Кауфмана для детей (*Kaufman Assessment Battery for Children, K* ABC*) (Kaufman & Kaufman, 1983). Этот тест дает оценку «умственной обработки информации» и состоит из различных более мелких тестов, или субтестов, на «последовательную обработку» и «одновременную обработку». Считается, что именно эти виды умственной активности являются основными у детей. Кроме того, учитывается особая оценка «достижений», отражающая уровень знаний, полученных в школе и за ее пределами.

Кауфманы оценили надежность своего теста несколькими способами. Например, они использовали процедуру, названную **расщепленная надежность.** Она заключается в том, что элементы, образующие конкретный субтест, разделяются пополам (например, четные и нечетные элементы), а затем находится корреляция между двумя половинами. Если тест надежный, то корреляция должна быть высокой — испытуемый, получивший высокую оценку для одной половины заданий, должен получить высокие оценки и для другой. Второй вид надежности называется **ретестовой надежностью** и отражает взаимосвязь между двумя различными случаями применения теста. Эта надежность также должна быть высокой — надежный тест дает непротиворечивые результаты от одной попытки тестирования к другой. В случае батареи тестов Кауфмана и расщепленная, и ретестовая надежности имеют высокие показатели. Но как обстоит дело с валидностыо теста?

Одним из показателей тестовой валидности является **критериальная валидность,** отражающая возможность на основании данного теста предсказать дальнейшие события. Этот вид валидности определяется путем исследования корреляции между изучаемым тестом (батарея Кауфмана) и оценками, полученными по некоторому «критерию». Критерием обычно является какой-либо другой тест или измерение, тематически связанное с интересующим тестом. В случае теста интеллекта в качестве критериальных оценок принимаются оценки школьной успеваемости, так как тесты интеллекта разрабатываются для того, чтобы иметь возможность предсказать академическую успеваемость. Оценки валидного теста должны положительно коррелировать с критериальными оценками.

Для проверки валидности батареи Кауфмана были проведены многочисленные исследования с использованием большого числа критериальных измерений достижений в учебе (Kaufman & Kaufman, 1983). В целом, результаты были впечатляющими: например, была выявлена корреляция батареи Кауфмана с Тестом базовых навыков штата Айова и Калифорнийским тестом достижений, официально принятыми в качестве тестов для оценки школьной успеваемости.

Этот пример говорит о важности использования хороших (т. е. надежных и валидных) тестов. Ведь на основании полученных результатов (по крайней мере отчасти) принимаются решения, способные повлиять на ход человеческой жизни. Дети обучаются по программе для одаренных, менеджеры продвигаются по службе, студенты колледжей поступают в медицинские вузы, а психиатры ставят своим пациентам правильные диагнозы. Все эти столь важные решения принимаются с помощью психологических тестов, а следовательно, надежность и валидность тестов имеет также и этическое значение. Во вставке 9.2 рассматриваются некоторые вопросы, посвященные этике тестирования, и приводятся нормы Американской психологической ассоциации по разработке и использованию тестов.

ВСТАВКА 9.2

Этика — нормы *APA* для психологического тестирования

Взгляните еще раз на табл. 2.2 и вы увидите, что переработанный в 1992 г. Этический кодекс *APA* содержит раздел, озаглавленный «Подготовка, оценивание или вмешательство». 10 стандартов, представленных в этом разделе, касаются разработки, использования и интерпретации психологических тестов. В частности, два стандарта связаны с вопросами исследований *(American Psychological Association,* 1992):

2.02. Компетентность, правильное оценивание и вмешательство.

а) Психологи, разрабатывающие, проводящие, оценивающие, интерпретирующие или использующие особые способы психологической оценки, интервью, тесты и другие методы, выполняют эту работу так, чтобы ее цель и способ выполнения соответствовали проводимому исследованию и чтобы существовало подтверждение корректности и правильного применения разрабатываемых методов.

2.03. Построение тестов.

Психологи, разрабатывающие и проводящие исследования, включающие тесты или другие методы оценки, используют научные процедуры и современные научные достижения по разработке, стандартизации, проверке эффективности, снижению или устранению искажений тестов, а также по разработке рекомендаций для их использования.

Рассмотрение этих стандартов, а также проведенное нами обсуждение надежности и валидности должны показать вам, что психологическое тестирование включает в себя гораздо больше, чем простое составление разумного опросника и проведение по нему опроса испытуемых. А именно так было разработано большинство псевдонаучных психологических тестов, встречающихся в популярных журналах. Подобные тесты кажутся научными, так как к ним прилагаются ключи для интерпретации оценок («если Ваши оценки лежат между 0 и 25, это означает, что...»), но предлагаемые шкалы не имеют никакого смысла, поскольку никто даже не пытался определить их надежность и валидность или провести стандартизацию, так чтобы можно было сравнить полученные оценки с данными, рассчитанными для всей популяции. Эти тесты вполне безвредны, если их не воспринимать всерьез. Вооружившись скепсисом, вы вполне можете поинтересоваться следующими вопросами:

- Испытываете ли Вы стресс? *(Self,* апрель 2000);
- Скучно ли с Вами? (*Young and Modern,* июнь/июль 1992);
- Проверь себя: злой или добрый? *(Self,* январь 1993);
- Ты и он: Ваши тайные сексуальные личности *(Glamour,* июнь 2000).

Чтобы научиться разрабатывать *настоящие* психологические тесты и проверять их эффективность, запишитесь на занятия по тестам и измерениям, которые наверняка проводятся на вашем факультете.

Пример 21. Исследования по психологии личности и анормальной психологии

При разработки тестов, измеряющих нормальные и патологические личностные черты, оценка надежности и валидности имеет огромное значение. Но корреляционные процедуры используются не только при разработке тестов, они также очень важны для исследований, посвященных изучению индивидуальных особенностей личности. К примеру, для проведения такого исследования можно отобрать большую группу людей, протестировать их на различные свойства личности, а затем

изучить корреляцию полученных оценок. Такое исследование может показать наличие положительной корреляции между интроверсией и тревожностью (интроверты склонны к тревожности) или отрицательной корреляции между интроверсией и социальностью (интроверты склонны избегать социальных контактов). В качестве примера можно привести исследование манеры объяснений и депрессии, проведенное Селигманом и его коллегами (Seligman et al., 1988).

В главе 3 вы познакомились с работой Селигмана, посвященной беспомощности. На ее примере была проиллюстрирована разработка исследований на основании теорий и развитие теорий под влиянием исследований. Селигман отталкивался от теории выученной беспомощности, основывающейся на нескольких исследованиях по выработке условных рефлексов у животных, а затем пересмотрел ее с учетом «объяснительного стиля». Согласно Селигману, страдающие депрессией люди долгое время попадали в ситуации, способствующие выработке у них приобретенной беспомощности, а их пессимистичная манера объяснять события собственной жизни коррелирует с депрессией. Исследование Селигмана (1988) было разработано для изучения взаимосвязи между пессимистическим стилем атрибуции и склонностью к депрессивным состояниям. Кроме того, исследователи хотели узнать, может ли «когнитивная терапия» иметь двойной эффект и, снижая депрессию, в то же время делать объяснительный стиль более позитивным.

Когда люди с пессимистическим объяснительным стилем переживают некоторое негативное событие, они обвиняют во всем себя, считают случившееся симптомом более общего персонального несоответствия действительности и думают, что будущие неудачи — это скорее правило, чем исключение (Abranson, Seligman & Teasdale, 1978). Вполне понятно, почему подобный образ мыслей сопровождается чувством депрессии. В исследовании Селигмана для выявления манеры объяснения использовался Опросник стиля атрибуции; уровень депрессии измерялся при помощи хорошо известного Опросника депрессии Бека. Исследователи обнаружили, что тяжесть депрессивного состояния значимо коррелирует с пессимистическим объяснительным стилем; данные подтвердились до проведения терапии ($r = +0{,}56$), после ее проведения ($r = +0{,}57$) и спустя 1 год после окончания терапии ($r = +0{,}63$). У пациентов по мере проведения терапии наблюдалось улучшение, но корреляция оставалась практически неизменной, так как «в ходе когнитивной терапии объяснительный стиль и депрессивное состояние изменялись одновременно и по мере того, как объяснительный стиль становился более оптимистичным, у пациентов снижался уровень депрессии» (Seligman et al, 1988, p. 17). Поскольку после терапии оценки объяснительного стиля и депрессии взаимосвязано изменились, то в результате корреляция практически не изменилась. Когнитивная терапия призвана изменить образ мысли людей, а следовательно, и то, как они объясняют происходящие с ними события, поэтому, по мнению исследователей, их работа подтвердила идею о том, что для эффективного лечения депрессии необходимо заставить людей изменить способ мышления и подход к интерпретации жизненных событий.

Одна из самых известных в психологии серий исследований является еще одним примером использования корреляционного подхода для изучения личности — в данном случае для изучения «мотивации достижения». Прежде чем продолжать

изучение материала, обратитесь ко вставке 9.3 — в ней описывается классическое исследование в этом направлении. Ученые, которые провели это исследование, сделали вывод, что для государства было бы правильным следить за тем, чтобы герои детских произведений всегда добивались успеха.

ВСТАВКА 9.3

Классические исследования — успешное общество

Смогли бы вы предсказать уровень успешности целого общества, проанализировав сказки, которые его члены рассказывают детям? Да, если опираться на книгу психолога Дэвида Мак-Клелланда «Успешное общество» (McClelland, *The Achieving Society,* 1961). Книга демонстрирует необычайно амбициозную попытку распространить результаты психологического исследования успешности на исторические факты. Мак-Клелланд и его коллега Джон Эткинсон были пионерами изучения мотивации, необходимой, чтобы переносить испытания и достигать успеха (McClelland, Atkinson, Clark & Lowell, 1953). Ученые разработали несколько способов выявления мотиваций успешности, провели впечатляющее количество исследований корреляции успеха и способствующих его достижению условий (подсказка: давайте детям побольше возможностей проявлять самостоятельность), и построили теорию мотивации успеха (Atkinson & Feather, 1966). Вследствие изучения «взаимодействия между устойчивыми мотивами, характерными для данной личности, и влиянием ситуации теория мотивации успеха представляет собой шаг в сторону концептуальной интеграции двух психологических дисциплин» (Atkinson & Feather, 1966, p. 5), к которой в 1957 г. призвал Ли Кронбах в своей известной президентской речи.

Один из способов оценки необходимости добиться успеха, или «НДУ», - это использование Тематического апперцепционного теста (ТАТ), в ходе которого испытуемые рассматривают изображения, допускающие различные толкования, и описывают то, что они видят (Murray, 1943). Например, изображение мальчика, разглядывающего скрипку, может быть прочитано как рассказ о мальчике, мечтающем стать скрипачом. Испытуемый, написавший такой рассказ, получит более высокую оценку НДУ, чем тот, кто представит, что мальчик хочет взять скрипку и ударить ей сестру по голове. Предполагается, что такие истории отражают стоящие за ними важные для испытуемого мотивы.

Идея о том, что на основании историй можно сделать вывод о мотивации человека, навела Мак-Клелланда на мысль проверить, какую роль играют детские сказки, басни и мифы в формировании мотивов молодых людей. Если в сказках будут преобладать темы достижения успеха, то не поможет ли это детям сформировать идею о важности успеха в жизни? Возможно ли определить общий уровень НДУ в обществе на основании анализа детской литературы, народных сказок, музыки и игр? И если дети растут в культурном окружении, делающем особый акцент на успешности, то поможет ли это им добиться успеха, когда они вырастут?
Такие размышления привели Мак-Клелланда к проведению исследования, посвященного общественным условиям успешности. Он подверг детскую литературу тому же виду анализа, который использовался для рассказов ТАТ, затем различными способами определил степень экономического благополучия общества и нашел корреляцию этих двух факторов. Была обнаружена положительная корреляция: по мере увеличения числа «тем успеха» в детских сказках наблюдался рост успешности в обществе. Особенно интересен тот факт, что успешность в реальности запаздывала по отношению к высокому уровню успешности в литературе примерно на 50 лет - как раз столько времени требуется, чтобы дети, испытавшие воздействие

литературы, часто обращающейся к теме достижения успеха, достаточно выросли, чтобы высокий уровень их НДУ смог повлиять на общество.

Идеи Мак-Клелланда не остались без критики. В данном исследовании проблема направленности была решена практически так же, как это делается в перекрестных корреляционных исследованиях, но остался нерешенным вопрос, наиболее распространенный в корреляционных исследованиях: проблема третьей переменной. Взаимосвязь между детской литературой и позднейшим успехом действительно очень впечатляет, но исторические тенденции чрезвычайно сложны и зависят от бесчисленного количества факторов. Тем не менее, исследование Мак-Клелланда стало классическим примером попытки использовать найденные психологические законы для объяснения исторических событий.

Пример 22. Вопрос о природе и среде

Как вы узнали из вставки 9.1, сэра Фрэнсиса Гальтона сильно впечатлил тот факт, что гении имеют тенденцию рождаются в одной и той же семье. Изучение сходства между членами семьи стало одной из ведущих стратегии исследования влияния наследственности и окружающей среды на различные характеристики человека. Обычно у всех пар родственников, входящих в одну семью, измеряются определенные характеристики, а затем находится корреляция между ними. Факторы наследственности и окружающей среды оцениваются независимо друг от друга: для этого сравниваются пары, различающиеся генетически, с парами, имеющими разные условия жизни. Например, однояйцевые близнецы, растущие вместе, сравниваются с однояйцевыми близнецами, разделенными при рождении и воспитывающимися в разных условиях, — таким образом, генетический фактор остается постоянным, а фактор окружающей среды изменяется. Аналогичным образом можно поддерживать более или менее постоянным фактор окружения и изменять генетический фактор. Для этого необходимо сравнить воспитывающихся вместе однояйцевых близнецов (генетически одинаковых) с воспитывающимися вместе двуяйцевыми близнецами (имеющими генетические различия). Подобные исследования обычно показывают совместное влияние врожденных особенностей и воспитания. Именно это отображено в табл. 9.2, обобщающей результаты десятков исследований, посвященных изучению природы интеллекта (Bouchard & McGue, 1981). Чем выше корреляция, тем ближе друг к другу показатели интеллекта у изучаемых пар. Так, корреляция +0,86 у однояйцевых близнецов означает, что если один из близнецов имеет высокий показатель интеллекта, то *IQ* второго близнеца также будет высоким. Очевидно, что корреляция снижается при увеличении генетических различий — это свидетельствует о том, как важна наследственность для интеллекта. Как показывает различие значений корреляции между однояйцевыми близнецами, воспитывающимися вместе (+0,86) и по отдельности (+0,72), среда также играет важную роль.

Так исторически сложилось, что большинство исследований близнецов посвящено вопросам интеллекта. Однако не так давно был проведен ряд исследований, показавших наследуемость особенностей личности и темперамента (Bouchard, Lykken, McGue, Segal & Tellegen, 1990). Обнаружилось, что даже застенчивость имеет определенные генетические основания. Имд и др. (Emde et al, 1992) исследо-

вали у 200 пар 14-месячных однояйцевых и двуяйцевых близнецов (воспитывающихся вместе) ряд личностных и когнитивных показателей. Одной из переменных была застенчивость, которую измеряли различными способами. Во-первых, она входила в общую оценку «поведенческого торможения», получаемую на основании анализа видеозаписи поведения детей, реагирующих на появление в комнате незнакомого человека. Отсутствие контакта с незнакомцем и стремление быть поближе к маме давали высокую оценку. Кроме того, застенчивость определяли по реакции на визит исследователя в семью ребенка и по поведению ребенка при первом посещении лаборатории. Родители также заполняли опросник, содержащий шкалу застенчивости. Как и в других исследованиях с близнецами, рассматривались корреляции для однояйцевых и двуяйцевых близнецов. Из табл. 9.3 видно, что близнецы с большим генетическим сходством получили более высокие оценки корреляции, что подтверждает гипотезу о существовании генетической зависимости для застенчивости.

Таблица 9.2

Корреляция между коэффициентами интеллекта у пар родственников, различающихся либо генетически, либо по условиям жизни

	Количество исследований	Количество пар	
Однояйцевые близнецы, воспитывающиеся вместе	34	4672	0,86
Однояйцевые близнецы, воспитывающиеся по отдельности	3	65	0,72
Двуяйцевые близнецы, воспитывающиеся вместе	41	5546	0,60
Родные братья или сестры, воспитывающиеся вместе	69	26473	0,47
Родные братья или сестры, воспитывающиеся по отдельности	2	203	0,24
Двоюродные братья или сестры	4	1176	0,15

Примечание. Приведенные данные представляют собой взвешенные средине значения корреляции из исследований, описанных Бошаром и Мак-Гыо (Bouchard & McGue, 1981).

Таблица 9.3

Корреляция между оценками «застенчивости» для пар однояйцевых и двуяйцевых близнецов

Оценка застенчивости	Значение пирсонова *r*	
	однояйцевые близнецы	двуяйцевые близнецы
Поведенческое торможение	0,57	0,26
Наблюдаемая застенчивость	0,70	0,45
Опросник, заполненный родителями	0,38	0,03

Источник: Из Emde et al., 1992, табл. 2.

Многомерный анализ

Двумерный анализ направлен на исследование взаимосвязей .между любыми двумя переменными. С помощью **многомерного анализа** изучаются взаимосвязи более чем двух переменных (часто количество переменных заметно превышает две). До сих пор в данной главе мы рассматривали случаи двумерного анализа, за исключением примера с частичной корреляцией, в котором оценивалось влияние третьей переменной на взаимосвязь между двумя другими. Далее вы коротко познакомитесь с двумя распространенными процедурами многомерного анализа — множественной регрессией и факторным анализом.

Множественная регрессия

В случае простой регрессии рассматриваются две переменные: предикторная и критериальная. Если оценки SAT(Американскин школьный тест проверки способностей) коррелируют со средним баллом первокурсников, то их можно использовать для предсказания академической успеваемости. Однако, как вы знаете из личного опыта, такой феномен, как «успеваемость в колледже» не так прост, как кажется. Оценки теста *SAT* могут говорить о будущей высокой успеваемости, но как быть с такими факторами, как «мотивация», «высокие школьные оценки» или «уклонение от занятий физикой»?

Множественная регрессия помогает решить проблему использования более чем одной предикторной нeременной. В исследовании, проводимом методом множественной регрессии, применяется одна критериальная, а также две или более предикторных переменных. Такой анализ позволяет вам не только выяснить, что на основании этих двух или более переменных можно предсказать определенный критерий, но также определить относительную предсказательную силу этих переменных. Эта сила отображается в формуле множественной регрессии для исходных данных, которая представляет собой расширенный вариант формулы простой регрессии:

Простая регрессия: $Y = a + bX.$

Множественная регрессия: $Y = a + b_1X_1 + b_2X_2 + ... + b_nX_n,$

где каждый X — это отдельная предикторная переменная, Y — это критериальная переменная, а величина показателей b отражает относительную важность каждой предикторной переменной. Этот показатель называют «весом регрессии» (Licht, 1995). В результате анализа по методу множественной регрессии получают множественный коэффициент корреляции *(R)* и множественный коэффициент детерминации *(R^2)*. R — это корреляция между объединенными предикторными переменными и критерием, а R^2 — показатель степени изменчивости критериальной переменной, вызванной объединенными предикторными переменными. Обратите внимание на использование больших букв для обозначения многомерных R *и* R^2 — тем самым их можно отличить от двумерных пирсонова *r и* r^2. Однако их интерпретации весьма схожи. И R, и r обозначают силу корреляции, а R^2 и r^2 — долю изменчивости, общей для нескольких переменных.

Преимущество анализа методом множественной регрессии заключается в том, что при объединении влияния нескольких предикторных переменных (особенно если они не слишком сильно коррелируют друг с другом) возможность сделать верные предположения заметно увеличивается в сравнении с обычным регрессионным анализом. Высокие школьные оценки уже сами по себе говорят о будущей хорошей успеваемости в колледже, так же как и оценки теста *SAT,* но на основании двух этих показателей можно делать предсказания с большей уверенностью (Sprinthall, 2000). Чтобы получить представление о том, в каких исследованиях применяют анализ методом множественной регрессии, рассмотрим следующие примеры.

1. Исследование, предсказывающее развитие эмпатии (сопереживания) на основании двух аспектов событий раннего детства (Barnett & McCoy, 1989): студенты, сопереживающие другим людям, обычно имели в детстве стрессовые переживания, что сделало их более чувствительными к переживаниям других. Тяжесть полученной в детстве травмы, как признак развития будущей эмпатии, имеет больший вес, чем общее число травматических событий.
2. Исследование, предсказывающее развитие восприимчивости к простуде на основании негативных событий жизни, воспринимаемого стресса и отрицательных эмоций (Cohen, Tyrell & Smith, 1993). Вы, возможно, думаете, что простуда развивается потому, что вы провели обеденное время слишком близко от непрерывно чихавшего человека. Однако данное исследование показало, что простудные заболевания можно предсказать на основании трех факторов, связанных со стрессом. Чаще всего простужаются те студенты, которые: а) в недавнее время пережили стрессовое событие, б) чувствуют, что к ним предъявляются завышенные требования, и в) описывают свой общий эмоциональный фон как негативный.

Факторный анализ

Вторая многомерная процедура — это **факторный анализ.** В ходе факторного анализа определяются значения большого количества переменных, находится корреляция между ними, а затем выявляется группы переменных, образующие «факторы». Поясним эту идею на простом примере. Предположим, вы дали школьникам следующие задания:

словарный тест (СЛ);

тест на понимание прочитанного (ПП);

тест на аналогии (например, доктор связан **с пациентом, как адвокат с_) (АН);**

тест по геометрии (ГЕОМ);

тест на решение головоломок (РГ);

тест на вращение фигур (ВФ).

Для всех возможных пар тестов можно **вычислить пирсоново *r*, в результате** получится так называемая **матрица корреляции:**

	СЛ	ПП	АН	ГЕОМ	РГ	ВФ
СЛ	–	+0,76	+0,65	−0,09	+0,02	−0,08
ПП		–	+0,55	+0,04	+0,01	−0,02
АН			–	−0,07	−0,08	+0,09
ГЕОМ				–	+0,78	+0,49
РГ					–	+0,68
ВФ						–

Обратите внимание на то, как некоторые значения корреляции образуют группы (я обвел две группы). Все корреляции между словарем, пониманием прочитанного и аналогиями довольно высоки. Это верно и для геометрии, головоломок и вращения фигур. Корреляции между тестами, принадлежащими к разным группам, практически равны нулю. Это говорит о том, что эти тесты направлены на исследование двух существенно различающихся умственных способностей, или «факторов». Мы можем обозначить их как «беглось речи» и «пространственные навыки».

Факторный анализ — это сложный статистический метод, с помощью которого из набора взаимных корреляций выделяются отдельные факторы. При анализе данной матрицы без сомнения будут выделены те же два фактора. В ходе анализа также определяются «факторные нагрузки», представляющие собой корреляции между каждым из тестов и каждым из выделенных факторов. В приведенном выше примере первые три теста будут иметь «высокую нагрузку» на фактор 1 (беглость речи), а вторые три — «высокую нагрузку» на фактор 2 (пространственные навыки). Конечно, в реальном исследовании корреляции никогда не группируются так четко, как в данном примере, и полученные результаты нередко приводят исследователей к бурным дискуссиям по поводу того, действительно ли обнаружены различные факторы. Также возникают расхождения в том, как правильно называть факторы, ведь факторный анализ сам по себе лишь выявляет факторы, а как их назвать — это решают сами исследователи.

Факторный анализ применялся в одном из самых долгих психологических споров — является ли интеллект единым свойством человека. Чарльз Спирмен — основоположник факторного анализа (начало XX в.), — считал, что все тесты интеллекта имеют одинаковую нагрузку на один фактор, который он назвал фактором общего интеллекта, или g (от англ. *general).* Более того, по его мнению, каждый тест должен давать высокую нагрузку на второй фактор, включающий навык, проверяемый данным тестом (например, математические способности). Эти факторы второго порядка, или «особые», он обозначил как s (от англ. *special).* Согласно его «двухфакторной» теории, выполнение тестов интеллекта напрямую зависит от общего интеллекта человека (g) и его особых навыков (л). Спирмен считал, что g наследуется, а различные 5-факторы приобретаются в процессе обучения (Fruchter, 1954).

Другие исследователи, и в том числе Льюис Терстоун, считали, что интеллект состоит из множества факторов, и отвергали существование общего фактора g. По результатам факторного анализа Терстоун сделал вывод, что существуют семь различных факторов, которые он назвал «первичными умственными способностями»

(Thurstone, 1938): понимание речи, беглость речи, навыки счета, пространственные навыки, память, скорость восприятия и способность к рассуждениям.

Вопрос о том, является ли интеллект единым целым, продолжает ставить в тупик ученых, занимающихся его измерением, и его обсуждение не входит в задачи этой главы. Для нас важно, что факторный анализ может привести к различным результатам. Это связано с тем, что а) существует несколько разновидностей факторного анализа, по-разному оценивающих, насколько высокой должна быть корреляция для выявления отдельных факторов, и б) в различных исследованиях этой проблемы используются различные тесты интеллекта. Поэтому исследователи, использующие различные подходы и тесты, получают самые разные результаты. Говоря коротко, так же, как и остальные статистические методы, факторный анализ — это лишь инструмент, и он не может сам по себе решать такие теоретические вопросы, как природа интеллекта.

Как стало ясно из этого небольшого введения, корреляционные процедуры играют заметную роль в современных психологических исследованиях. Очень часто в них возникает необходимость, если экспериментальные процедуры использовать невозможно. Кроме того, разработка сложных многомерных процедур упростила решение вопроса о причинах и следствиях по сравнению с прошлым, когда большинство корреляционных процедур были двумерными по своей природе.

Многие корреляционные исследования проходят за пределами лабораторий. В следующей главе мы более подробно изучим вопрос прикладных исследований и рассмотрим в деталях несколько так называемых квазиэкспериментальных планов. Также будет рассмотрен особый случай прикладных исследований — исследования, посвященные оценке программ. Они приобретают все большее значение для системы социального обеспечения и образования.

Резюме

Две психологические дисциплины

Так же как и экспериментальное исследование, корреляционное исследование — это один из двух традиционных научных психологических подходов. В ходе эксперимента производится непосредственное управление переменными и наблюдаются результаты, а в процессе корреляционного исследования наблюдаются взаимосвязи между естественными переменными.

Корреляция и регрессия — основы

Две переменные коррелируют, если между ними существует действительная взаимосвязь. Взаимосвязь является прямой в случае положительной корреляции и обратной, если корреляция отрицательная. Силу взаимосвязи можно оценить по графику рассеяния, по абсолютному значению коэффициента корреляции (например, по пирсонову r*,* если данные получены по интервальной шкале или шкале отношений) или по коэффициенту детерминации, который определяется возведением в квадрат коэффициента корреляции. Если известна корреляция, то с помощью регрессионного анализа можно сделать предположения о будущих событиях. Если между двумя переменными *A b B* наблюдается значимая корреляция, то, зная значение *A,* можно с определенной вероятностью предсказать значение *B.*

Интерпретация корреляций

Наличие значимой корреляции между переменными *А* и *В* само по себе не позволяет сделать вывод о том, что *А* является причиной *В*. Проблема направленности заключается в том, что причинно-следственная связь может проходить в любом из двух направлений: *А* вызывает *В* или *В* вызывает Л. Проблема направленности уменьшается, если между измерениями *А* и *В* проходит некоторое время. Проблема третьей переменной состоит в том, что для многих корреляций взаимосвязь между неременными является следствием совместного влияния неконтролируемых переменных, естественным образом ковариирующих с изучаемой переменной. Это значит, что некоторая третья переменная может вызвать изменения переменных Л и *В*. Влияние третьей переменной можно оценить с помощью процедуры взаимной корреляции.

Использование корреляций

Необходимость в корреляционных исследованиях часто возникает, если по практическим или этическим причинам нельзя провести эксперимент. Этот подход часто используется в исследованиях по оценке психологических тестов (т. е. при определении их надежности и валидности), в исследованиях, посвященных изучению индивидуальных различий в области психологии личности и психопатологии, а также в исследованиях близнецов и при изучении вопросов, связанных с ролью наследственности и окружающей среды в деле формирования определенных свойств.

Многомерный анализ

Двумерный анализ призван исследовать взаимосвязь между двумя переменными. Многомерный анализ посвящен изучению перекрестных взаимосвязей между более чем двумя переменными. С помощью множественной регрессии на основании двух или более предикторных переменных можно предсказать получение определенного результата (критерия). Факторный анализ выявляет группы факторов, лежащие в основе большого числа взаимосвязей.

Задания для повторения

Выбор ответа

1. При каком из следующих значений корреляции взаимосвязь наиболее сильная?
   1) +0,81;
   2) -0,67;
   3) -0,86;
   4) +1,04.
2. В корреляционных исследованиях внешние переменные не контролируются, что приводит к проблеме интерпретации, известной как:
   1) проблема направленности;
   2) проблема регрессии к среднему;
   3) проблема ограничения диапазона;
   4) проблема третьей переменной.

3. Руководство аспирантуры стало принимать решения о приеме кандидатов с помощью определенной формулы. В формулу входят оценки кандидата за 4 года обучения, оценки по профилирующему предмету и оценки *GRE* (американский выпускной тест). Оценки *GRE* имеют наибольший вес; за ними идут оценки по профилирующему предмету. Какая процедура была использована для создания такой формулы?
   1) множественная регрессия;
   2) двумерная регрессия;
   3) факторный анализ;
   4) взаимная корреляция.
4. В XIX в. большинство людей считали, что интеллект можно оценить, измерив объем мозга. Какое из следующих высказываний по поводу такого подхода верно?
   1) он обладает критериальной валидностью;
   2) он пройдет проверку на ретестовую надежность;
   3) он не будет ни надежным, ни валидным;
   4) он будет валидным, но не будет надежным.
5. Перекрестный метод иногда используется для решения проблемы:
   1) третьей переменной;
   2) регрессии к среднему;
   3) отрицательной корреляции;
   4) направленности.

Короткие эссе

1. В чем заключаются основные идеи речи Ли Кронбаха о «двух дисциплинах»?
2. Опишите, чем отличаются графики рассеяния для а) положительной и отрицательной корреляции и б) сильной и слабой корреляции.
3. Что такое коэффициент детерминации и о чем говорит его значение? Для иллюстрации ответа используйте пример с депрессией и средним баллом.
4. Исследователь обнаружил, что люди с высокой самооценкой обычно больше занимаются спортом, чем те, кто имеет низкую самооценку. Объясните, в чем заключается проблема направленности и как она влияет на интерпретацию подобной корреляции.
5. Исследователь обнаружил, что дети, много играющие в видеоигры, склонны к проявлениям агрессии против своих сверстников в школе. Объясните, в чем заключается проблема третьей переменной и как она влияет на интерпретацию подобной корреляции.
6. На примере исследования Эрона и др. (Eron et al. 1972), в котором была обнаружена значимая взаимосвязь между телевизионными предпочтениями в 3-м классе и агрессией в 13-м классе, покажите причину, по которой используется процедура частичной корреляции.

7. Дайте определение надежности и опишите два способа ее получения.
8. На примере исследования, связанного с тестом Кауфмана K· ABC, покажите значение критериальной валпдности.
9. В результате широко известного исследования, посвященного историческому рассмотрению мотивации успеха, Дэвид Мак-Клеллапд пришел к выводу, что общество должно стремиться развивать самостоятельность детей и тем самым делать их способными добиваться успеха. Опишите метод, позволивший Мак-Клелланду сделать такой вывод.
10. Опишите, как использовать корреляционные процедуры для изучения вопроса природы и среды.
11. Охарактеризуйте отличия двумерного и многомерного методов анализа и опишите процедуру множественной регрессии.
12. Опишите основные этапы факторного анализа.

Упражнения

Упражнение 9.1. Интерпретация корреляций

Ниже описаны результаты гипотетических двумерных корреляционных исследований. Проинтерпретируйте результаты по крайней мере двумя способами с учетом проблем направленности и третьей переменной.

1. Существует положительная корреляция между уровнем авторитарности матери и застенчивостью ребенка.
2. Существует отрицательная корреляция между депрессией и уровнем физической подготовки.
3. Существует положительная корреляция между объемом домашней библиотеки и средним баллом студента в колледже.
4. Счастливые в браке пары обычно имеют больше сексуальных контактов (друг с другом), чем несчастливые.
5. Существует отрицательная корреляция между оценками и боязнью проходить тестирование.
6. Место, занимаемое студентом в классе, коррелирует с его оценками — чем ближе к преподавателю сидит студент, тем выше его оценки.

Упражнение 9.2. График рассеяния, вычисление пирсонова r и регрессия

По приведенным ниже данным постройте график рассеяния. На основании графика предположите, какой должна быть корреляция, а затем с помощью процедуры, описанной в приложении *С,* вычислите действительное значение пирсонова r. Затем найдите коэффициент детерминации и опишите обнаруженную взаимосвязь. В заключение подставьте полученные данные в уравнение регрессии и сделайте предположение о «манере объяснять» Эда (его значение самооценки равно 10), Неда (самооценка равна 20) и Фреда (самооценка равна 30).

1. Переменная *A* = самооценка: оценки варьируются от 0 до 50; более высокие оценки говорят о более высокой самооценке.
2. Переменная *B* = «манера объяснять»: оценки варьируются от 0 до 75; более высокие оценки говорят об отрицательном, или пессимистичном, способе интерпретации жизненных сложностей, а более низкие оценки — о положительном, или оптимистичном, способе объяснения тех же событий.

N испытуемого	Переменная A	Переменная B
1	42	32
2	22	34
3	16	65
4	4	73
5	46	10
6	32	28
7	40	29
8	12	57
9	28	50
10	8	40
11	20	50
12	36	40

Упражнение 9.3. Интерпретация графиков рассеяния

Постройте графики рассеяния, приблизительно отображающие нижеследующие взаимосвязи. Напишите отчет о каждой взаимосвязи.

1. Корреляция со значением +0,50 между субтестами на последовательную и одновременную обработку информации из теста Кауфмана К» ABC.
2. Корреляция со значением -0,80 между средним баллом и оценками теста «телесериалыюго *IQ*».
3. Корреляция со значением -0,02 между интеллектом и депрессией.
4. Корреляция со значением +0,90 между склонностью к неврозу навязчивых состояний и любовью к экспериментальной психологии.

ГЛАВА 10

Квазиэкспериментальные планы и прикладные исследования

Обзор задач главы

В этой главе мы будем рассматривать прикладные исследования, и в частности исследования по оценке программ. С прикладными исследованиями вы встречались в главе 3 при сравнении их с фундаментальными исследованиями. Из этой главы вы узнаете, что прикладные исследования представляют собой мощное направление в американской экспериментальной психологии и отображают один из центральных моментов американской идеологии — идею полезности. Оценка программы, призванной определенным образом помочь людям, — это вид прикладного исследования, направленный на изучение ее эффективности. В данной главе также описываются так называемые квазиэкспериментальные планы, часто используемые в прикладных исследованиях, в том числе в исследованиях по оценке программ[1]. Изучив данную главу:

- вы сможете определить двойное назначение прикладных исследований;
- поймете, почему прикладная психология всегда являлась важной частью американской психологии;
- сумеете охарактеризовать планы и этические проблемы, связанные с прикладными исследованиями, и поймете, какие из планов, описанных в предыдущих главах, были квазиэкспериментальными;
- сможете описать особенности плана с неэквивалентными контрольными группами и поймете, почему такой план обязательно содержит осложнители;
- поймете, почему уравнивание неэквивалентных групп по оценкам предварительного тестирования может вызвать эффект регрессии;
- научитесь описывать особенности планов с прерванной временной последовательностью и поймете, как с их помощью можно оценивать существующие тенденции;

[1] Очень важно обратить внимание па то, что совместное рассмотрение квазиэкспериментальных и прикладных исследований, осуществляемое в этой главе, не означает, что квазиэксперпментальные планы встречаются только в прикладных исследованиях. Чаще всего их можно встретить именно в таких исследованиях, но фундаментальные исследования также могут проводиться с использованием названных планов. Более того, их совместное рассмотрение не означает, что в прикладных исследованиях нельзя использовать экспериментальные планы.

- сумеете описать стратегии проведения анализа потребностей;
- поймете цели и процедуру проведения предварительной оценки, суммарной оценки и оценки затрат-эффективности;
- сможете охарактеризовать особые этические проблемы, часто сопутствующие исследованиям по оценке программ.

Как было отмечено в самом начале, главная моя цель —развить в вас интерес к изучению законов поведения и желание стать психологом-экспериментатором. Мой преподавательский опыт подсказывает мне, что некоторые из вас действительно будут участвовать в проведении исследований, но большинство не станет заниматься этим. Тем не менее многие из вас станут профессиональными психологами и посвятят себя работе консультанта, специалиста по работе с персоналом или другой психологической или связанной с охраной здоровья деятельности (например, физиотерапии). Поэтому вы встретитесь в своей работе со сферой прикладных исследований и оценки программ. Очень возможно, что вам придется:

- читать, понимать и критически оценивать литературу, посвященную исследованиям по оценке программ, которая будет использоваться службой, в которой вы работаете;
- участвовать в планировании новых программ, информируя (тактично) тех, кто незнаком с техникой исследований, об адекватности их оценки программ;
- участвовать в усовершенствовании методов и подходов, используемых в вашем учреждении для получения аккредитаций. Возможно, это будет связано с вашими знаниями по курсу методов исследований;
- разрабатывать исследование по оценке программы какого-либо учреждения и нести за него ответственность.

Другими словами, эта глава важна вне зависимости от того собираетесь ли вы когда-либо проводить психологические исследования.

За пределами лаборатории

Впервые вы узнали о различии между фундаментальными и прикладными исследованиями из главы 3. Важнейшей целью фундаментальных исследований по психологии является расширение наших знаний о законах поведения. Эти знания могут затем найти применение, но не это является основным мотивом работы — знание ценится само по себе. Прикладные исследования призваны увеличить наши знания о причинах существующих проблем и нацелены па их решение. Второе различие заключается в том, что фундаментальные исследования обычно проводятся в стенах лаборатории, находящейся в университетском городке или в кампусе колледжа, а прикладные исследования чаще всего проходят вне академической обстановки, в клиниках, учреждениях социального обеспечения, тюрьмах и на предприятиях. Конечно, существуют и исключения: некоторые фундаментальные исследования проводятся в «полевых» условиях, а некоторые прикладные — в лабораториях.

Чтобы дать вам представление о разнообразии прикладных исследований, рассмотрим заголовки нескольких статей, опубликованных в двух видных психоло-

гических журналах, «Журнале прикладной психологии» и «Журнале экспериментальной психологии»:

«Пол и взаимосвязи между оценкой справедливости оплаты труда или между продвижением по службе и удовлетворением от работы» (Witt & Nye, 1992);
«Зрительные пространственные способности пилотов» (Dror, Kosslyn, & Waag, 1993);
«Осознавание обстановки, сопутствующей вождению автомобиля: эксплицитное и имплицитное знание в динамической пространственной памяти» (Gugerty, 1997);
«Явление вероятности и практические аспекты программ по сохранению видов, находящихся под угрозой вымирания» (DeKay & McClelland, 1996).

Эти заголовки отображают две особенности прикладных исследований. Во-первых, в соответствии с приведенным выше определением данные исследования фокусируются на легко локализуемых проблемах: от удовлетворения работой до охраны вымирающих видов. Во-вторых, такие названия показывают, что хотя главной целью прикладных исследований является решение конкретных проблем (например, улучшение навыков вождения), они также способствуют углублению наших знаний о фундаментальных психологических процессах (например, о пространственной памяти). И действительно, как показывает исследование, разработанное для оценки так называемой «теории справедливости» (Adams, 1965), между фундаментальными и прикладными исследованиями существует тесная связь. Основное допущение названной теории состоит в том, что для оценки взаимоотношений с другими людьми мы сравниваем относительный вклад, сделанный нами в эти взаимоотношения, с относительной полученной пользой. Если затраты не соответствуют полученной пользе, мы ощущаем ситуацию как несправедливую и стремимся восстановить баланс. Например, если в супружеской паре оба супруга работают, а домашним хозяйством занимается в основном жена, то она может почувствовать несправедливость и попытаться прибегнуть к различным процедурам (например, договориться с мужем), чтобы нормализовать положение.

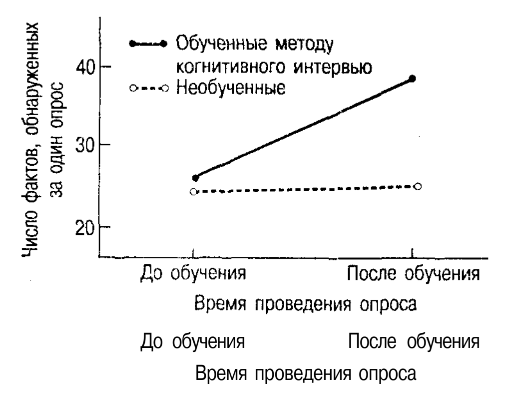
Большинство прикладных исследований выполняют двойную функцию: они одновременно решают прикладные задачи и подтверждают фундаментальные психологические явления, тем самым способствуя разработке теорий. Справедливость, к примеру, часто изучается именно прикладными исследованиями. Указанное выше исследование корреляции между полом и оценкой оплаты труда или продвижения по службе затрагивает важную производственную проблему (как увеличить удовлетворение от работы), но также дает информацию, касающуюся более общего вопроса адекватности теории справедливости. Пример, иллюстрирующий оценку метода развития способности к опознанию, также показывает эту особенность прикладных исследований.

Пример 23. Прикладное исследование

Вспомните исследование с участием шотландского клуба ныряльщиков (глава 8), показавшее, что запомненная информация воспроизводится лучше всего, если и запоминание, и воспроизведение происходят в одном и том же месте (например, под водой). Важность контекста для запоминания и общеизвестная неточность свидетельств очевидцев навели Эдварда Гейзельмана и Рональда Фишера на мысль

о разработке особого подхода, который они назвали «когнитивным интервью». Одна из особенностей этого подхода — принцип, обозначенный исследователями как «соответствие воспроизводимого контекста обстоятельствам события *{event-intemiew similarity)»,* применяя который исследователь старается «восстановить в памяти свидетелей внешние (например, погода), эмоциональные (чувство страха) и когнитивные (мысли) аспекты, переживаемые в момент преступления» (Fisher, Geiselman, & Amador, 1989, p. 723). Говоря кратко, исследователь пытается помочь свидетелю мысленно воспроизвести контекст произошедших событий.

Гейзельман и его коллеги ранее продемонстрировали эффективность своего подхода в условиях лаборатории, но они хотели оценить метод «за пределами знакомых стен» (Fisher et al., 1989, p. 742). Они обучили методу когнитивного интервью семь полицейских из отдела ограблений полиции Майами штата Флорида. В сравнении с девятью полицейскими из контрольной группы обученные полицейские при опросе свидетелей получали более надежную информацию (т. е. подтверждаемую другими источниками). Из рис. 10.1 видно, что обе группы показали приблизительно одинаковые результаты на этапе предварительного тестирования, включавшем 4 месяца, на протяжении которых полицейские записывали свои опросы свидетелей, а независимые судьи оценивали полученные сведения. После того как экспериментальная группа завершила обучение, обе группы провели и записали несколько опросов, оцененных таким же образом. Выяснилось, что обученные методу когнитивного интервью полицейские способны получить от свидетелей больше информации.



До обучения После обучения
Время проведения опроса

Рис. 10.1. Проверка эффективности метода когнитивного интервью 8 ходе прикладного исследования (из работы Фишера и др., 1989)

Исследование Фишера и др. — это превосходный пример того, как прикладное исследование помогает решить конкретные проблемы, а также способствует расширению наших знании о фундаментальных психологических явлениях. Оно показывает, как особый метод опроса может увеличить эффективность работы полиции, и в то же время дает новые подтверждения важности контекста для запоминания.

Прикладная психология в историческом контексте

Поскольку психология в Америке формировалась в академических условиях, вы можете решить, что она традиционно развивалась в направлении фундаментальных исследований. Это не так. С тех пор, когда в конце XIX в. сформировалась

новая наука, психологи в Соединенных Штатах интересуются прикладными исследованиями. Во-первых, под давлением общества в начале XX в. психологам пришлось показать, что они способны сделать что-то для изменения общества к лучшему. Чтобы получить приличный кусок от пирога академического финансирования в те времена, когда психологические лаборатории только открывались, психологи должны были продемонстрировать, что их идеи могут найти себе применение.

Психологи-исследователи занимались не только расширением знаний, но и нередко старались использовать методы фундаментальных исследований для решения вопросов образования, психического здоровья, воспитания детей и, как в случае Уолтера Майлза, спорта. Майлз в 1920 г. был директором лаборатории при Стэнфордском университете. Он посвятил свою работу в основном фундаментальным исследованиям и тем не менее в 1927 г. оказался на футбольном поле, где практиковалась университетская футбольная команда (рис. 10.2, Майлз в пиджаке). Легендарный стэнфордский тренер «Поп» Уорнер («Pop» Warner) был известен своим новаторским подходом к тренировкам и интересовался всеми методами, способными улучшить игру команды. Появляется Майлз, разработавший «сложный хронограф» для одновременного определения скорости реакции у семи футболистов (Miles, 1928). По сигналу семь мячей для гольфа падали во вращающийся с постоянной скоростью барабан (по одному мячу на каждого игрока) и семь игроков выполняли рывок вперед, толкая раму, которая тянула за нить, выбрасывающую в барабан еще семь мячей. Мячи оставляли отметки на барабане и, зная скорость его вращения, а также расстояние между отметками, Майлз определял скорость реакции игроков. Ученый провел серию исследований со сложным хронографом (см., например, Miles, 1931) и показал, что его можно использовать для отбора игроков с наивысшей скоростью реакции, но прибор не получил широкого применения.

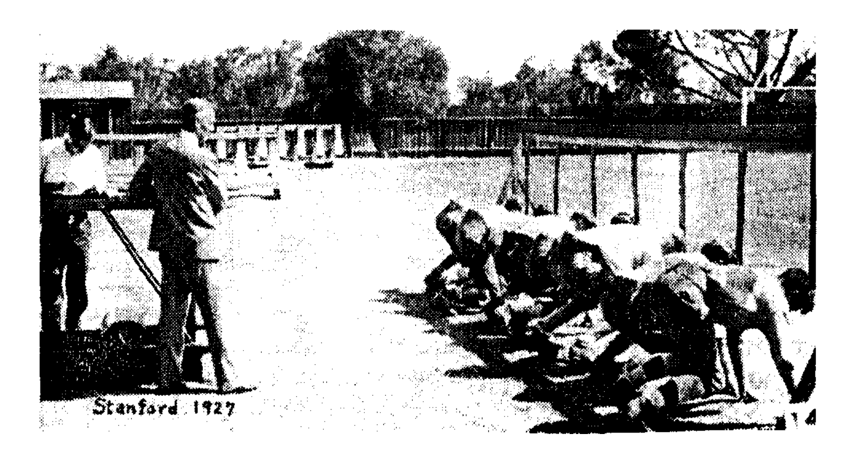


Рис. 10.2. Одновременная проверка скорости реакции у членоз стэнфордской футбольной команды, около 1927 г. (из архивов по истории американской психологии, Университет Экрон, город Экрон, штат Огайо)

Более основательная попытка применения психологических принципов была произведена в первой клинике для коррекции психических проблем, основанной Лайтнером Уитмером при Университете штата Пенсильвания. Уитмер учился на психолога-экспериментатора в лаборатории Вундта в Лейпциге и с 1890 г. стал активно заниматься экспериментами. Однажды (в 1895 г.) он выступил за отделение психологов-экспериментаторов от недавно образовавшейся Американской психологической ассоциации и за образование отдельной организации, так как видел недостаточную научную подготовку *APA* (Goodwin, 1985). В марте 1896 г. школьный учитель привел в лабораторию Унтмера маленького мальчика с необучаемостью (так мы называем подобное явление сегодня), и это событие отметило основание первой американской психологической клиники (McReynolds, 1987). Уитмер посвятил свою работу применению всех своих знаний о психических процессах для лечения различных психологических дисфункций, а в особенности нарушений способности к обучению, влияющих на школьную успеваемость.

Интерес психологов к прикладной области хорошо виден на примере движения по умственному тестированию. В предыдущей главе мы говорили о том, что тестирование интеллектуальных и личностных особенностей играет в Америке очень большую роль. Корни этого направления можно найти в работах первых психологов. Например, Джеймс Кеттелл, так же как и Уитмер, выпускник лаборатории Вундта, в 1890 г. ввел в использование термин «умственный тест» и в 1890-х гг. с учетом открытий Гальтона, сделанных в Англии, разработал в Колумбии расширенную (но обладавшую непоправимыми изъянами — см. Sokal, 1987) программу тестирования. В начале 20-х гг. XX в. он основал существующую и сегодня Психологическую корпорацию — гиганта индустрии психологического тестирования (например, она издает все тесты интеллекта Векслера).

И в завершение рассмотрим научную жизнь Гуго Мюнстерберга — идеальный для Америки пример прикладного психолога. Как и многие другие первые психологи, он учился на психолога-исследователя, но стал известен как эксперт по применению психологических подходов в бизнесе, клинической психологии и даже в суде. Одно время он был одиозной фигурой в Америке, и поэтому до недавних пор при историческом рассмотрении прикладной психологии его имя обычно опускалось. Чтобы более подробно узнать о человеке, заслужившем титул первого настоящего прикладного психолога Америки (несмотря на то, что он был немец), прочитайте вставку 10.1.

В начале XXI в. психологи заинтересованы в практическом применении своих открытий так же сильно, как и их предшественники в XX в. Профессиональные психологи не просто применяют научные знания о поведении для решения проблем, с которыми они сталкиваются. Они также выполняют исследования, соответствующие приведенному выше определению прикладного исследования. Это означает, что они разрабатывают и проводят исследования, чтобы найти решение конкретных проблем и одновременно чтобы обогатить фундаментальную базу психологических знаний. Однако прикладным исследованиям свойственны различные трудности, которые не встречаются в лабораториях.

ВСТАВКА 10.1

История — Гуго Мюнстерберг, экстраординарный психолог

Неудивительно, что прикладная психология представляет собой центральную силу американской психологии. В конце концов американцы славятся своим неизменным прагматизмом. «Какая в этом польза?» - такой вопрос является для нас вполне естественным. Примечательно поэтому, что одним из самых выдающихся прикладных психологов Америки считается Гуго Мюнстерберг (1863—1916), немец по национальности. Когда в 1892 г. он приехал в страну с целью основать психологическую лабораторию в Гарварде, он едва говорил по-английски. К моменту своей смерти он написал около 20 английских книг и был признан пионером производственной и судебной психологии и ученым, внесшим заметный вклад в клиническую психологию.

«Психологию и эффективность производства» Мюнстерберга, опубликованную в 1913 г., иногда считают первой вехой в развитии производственной психологии (Goodwin, 1999). Она включает в себя советы и описания исследований на такие темы, как подбор людей для определенных видов работ, факторы, наиболее сильно влияющие на производительность труда, а также предложения по маркетингу и рекламе. В работу также вошли его собственные исследования по разработке тестов для подбора персонала, в которых использовались процедуры проверки валидности, ставшие сегодня стандартными. К примеру, когда Мюнстерберга попросили разработать систему отбора для Новой английской телефонной компании, он проанализировал работу операторов с точки зрения решения познавательных задач (он выделил 14 видов задач), разработал тесты для проверки способности к решению каждого из видов и подтвердил валидность тестов, показав, что люди, получающие наивысшие оценки, становятся лучшими операторами.

Мюнстерберг был также первопроходцем в области судебной психологии - применении психологических принципов в юриспруденции. Его книга «Сточки зрения свидетеля», вышедшая в 1908 г., содержала поразительное описание того, как опасно верить свидетельским показаниям, в ней он высказывал предостережения, которые позднее были сформулированы в XX в. в ходе экспериментальной работы над этим вопросом. Не имея непосредственного экспериментального подтверждения несовершенства памяти свидетелей, для того чтобы показать ненадежность наблюдений людей, Мюнстерберг использовал свои знания об иллюзиях восприятия. Он также утверждал, что различные психологические методы могут быть использованы, чтобы уличить во лжи. Однако, как это часто бывает, если ученый попадает в поле зрения широкой публики, деятельность Мюнстерберга была поставлена под сомнение. Ссылаясь на исследование, показавшее, что суждение группы более адекватно только в случае испытуемых-мужчин, Мюнстерберг заключил, что суд присяжных не должен включать женщин. Его вывод получил широкую огласку в прессе и навлек на него гнев публики, выраженный заголовком газетной статьи, процитированной в книге Хотерсэла (1990 г.): «Разгневанные Мюнстербергом суфражистки считают, что женщины вполне подходят для суда присяжных».

Еще большую полемику вызвала попытка Мюнстерберга нарисовать в благоприятном свете портрет своей родной Германии. К сожалению, он сделал это прямо перед началом Первой мировой войны, что привело к лавине пышущих ненавистью писем с обвинениями в том, что Мюнстерберг - немецкий шпион. Согласно газете Хотерсэлу (1990), некто даже предложил Гарварду 10 млн долларов за увольнение Мюнстерберга, но отказался от своего предложения, когда ученый согласился подать в отставку, если этот человек заплатит 5 млн университету и 5 млн самому Мюнстербергу! Когда ученый в 1916 г. перед самым окончанием войны умер, его считали одним из злодеев Америки. «Американский психологический журнал», обычно печатающий длинные некрологи о психологах гораздо меньшего значения, чем Мюнстерберг, даже не упомянул о его смерти. Все это привело к тому, что значение этого ученого как первопроходца и человека, всячески способствовавшего развитию прикладной психологии, было забыто на долгие годы (Landy, 1992).

Проблемы прикладных исследований

Поскольку вы изучили главы 2, 5 и 6, для вас не будет неожиданностью материал данного раздела. Прикладные исследования связаны со следующими проблемами.

- *Этические дилеммы* (глава 2). Ученые, проводящие полевые исследования, могут столкнуться с проблемами обоснованного согласия и ограждения личной жизни. Также не всегда возможно проведение дебрифинга. Исследования, проходящие в производственных условиях, могут содержать элемент принуждения, если работники решат, что их служебное положение зависит от участия в исследовании (более подробно об этике прикладных исследований см. вставку 10.3).
- *Компромисс между внешней и внутренней валидностью* (глава 5). Поскольку исследования по прикладной психологии часто проходят в полевых условиях, исследователь может отчасти потерять контроль за переменными, влияющими на ход исследования. Поэтому опасность возможного осложнения снижает внутреннюю валпдность исследования. С другой стороны, прикладные исследования обычно имеют высокую внешнюю валпдность, так как условия их проведения более близки к ситуациям реальной жизни и в их ходе решаются повседневные проблемы,
- *Проблемы, свойственные только межсубъектным планам* (глава 6). В прикладных исследованиях для формирования групп не всегда возможно использовать случайное распределение, поэтому нередко приходится сравнивать неэквивалентные группы. Это, конечно, может снизить внутреннюю ватидность исследования вследствие проблем отбора или наличия взаимодействия между отбором участников и другими факторами, такими как взросление. При попытке использовать уравнивание, чтобы добиться определенного уровня эквивалентности, может возникнуть проблема регрессии (глава 5).
- *Проблемы, свойственные только внутрисубъектным планам* (глава 6). В прикладных исследованиях с внутрисубъектными факторами не всегда возможно провести необходимое позиционное уравнивание. Поэтому в исследовании может возникнуть неконтролируемый эффект последовательности. Исследования, продолжающиеся длительный период, нередко также сталкиваются с проблемой истощения.

Прежде чем продолжить дальнейшее изучение главы, вам необходимо вновь обратиться к соответствующим разделам глав 2,5 и 6 и повторить материал, посвященный указанным идеям. Кроме того, вам стоит повторить раздел главы 5 о том, какие выводы могут быть сделаны в случае управляемых и субъективных переменных.

Квазиэкспериментальные планы

Строго говоря, «настоящие» экспериментальные исследования проводятся с использованием управляемых независимых переменных и либо эквивалентных групп в случае межсубъектных планов, либо позиционного уравнивания в случае внутрисубъектных планов. Все остальные исследования являются квазиэкспериментальыгыми («почти» экспериментальными). В главе 5 обсуждение неуправляемых переменных основывалось на том, что субъектные переменные используются в каче-

стве независимых. В данной главе обсуждение этой темы будет несколько расширено. В целом, о **квазиэксперименте** говорят тогда, когда вследствие неполного контроля за переменными нельзя сделать выводы о существовании причинно-следственной связи. Мы уже встречались с некоторыми квазиэкспериментальными планами:

однофакторный план с неэквивалентными группами, включающий один или несколько уровней;

факторный план с неэквивалентными группами;

факторный план Р х Е;

все корреляционные планы.

В этой главе мы рассмотрим два вида планов, которые обычно описываются в работах, посвященных квазиэкспериментальным планам (например, Cook & Campbell, 1979): планы с неэквивалентными группами и планы с прерванной временной последовательностью. Существуют и другие квазиэкспериментальные планы, но эти два встречаются наиболее часто.

Планы с неэквивалентными контрольными группами

Такие исследования проводятся для оценки эффективности воздействия программ. Люди, участвующие в программе, сравниваются с членами контрольной группы, **не** подвергающимися воздействию. Этот вид планов используется, когда нельзя применить случайное распределение. Кроме разницы в значениях независимой переменной контрольная группа имеет и другие отличия от группы воздействия. Это значит, что группы неэквивалентны с самого начала исследования. В плане такого исследования вы должны узнать особый случай, обозначенный в главе 7 как план с «неэквивалентными» группами, в котором сравниваются неидентичные группы, нередко отбираемые на основании таких субъектных переменных, как пол, возраст или свойства личности. В случае **плана с неэквивалентной контрольной группой** группы не равны друг другу с начала исследования, а *кроме того,* в ходе исследования они участвуют в разных событиях. Поэтому использование такого плана влечет за собой осложнение, способное затемнить интерпретацию исследования. Тем не менее подобный план может быть весьма полезен для оценки программы воздействия, когда невозможно использовать случайное распределение.

В соответствии со схемой, предложенной Кэмпбеллом и Стэнли (Campbell & Stanley, 1963), план с неэквивалентной контрольной группой можно представить следующим образом.

Экспериментальная группа:	H_1	B	H_2.
Неэквивалентная контрольная группа:	H_1		H_2,

где Hj и H_2 обозначают соответственно доэкспериментальные и послеэкспериментальные наблюдения или измерения, а В обозначает оцениваемую программу воздействия. Поскольку группы могут различаться на стадии предварительного тестирования, сравнение двух групп заключается не просто в изучении разницы результатов, полученных на этапе заключительного тестирования, а в сравнении степени изменений, произошедших между предварительным и заключительным тестированием. Таким образом, статистически сравниваются оценки изменений (разница между H_1 и H_2) для каждой группы. Рассмотрим это на конкретном примере.

Предположим, что руководство компании по производству электрических сковородок хочет ввести новый свободный график рабочего дня. Рабочие будут продолжать работать по 40 часов в неделю, но в соответствии с новым расписанием они смогут начинать и заканчивать рабочий день в разное время или отрабатывать все 40 часов за 4 дня и отдыхать 3 дня в неделю. Руководство надеется, что это нововведение поднимет рабочий настрой, а следовательно, увеличит производительность труда, и поэтому разрабатывает квазиэксперимент для проверки предположения. Компания владеет двумя фабриками: одной вблизи от Питтсбурга и второй неподалеку от Кливленда. Бросив монетку, руководители компании решили сделать из ппттсбургской фабрики экспериментальную группу, а из кливлендской — неэквивалентную контрольную группу. Очевидно, что квазиэкспериментальное исследование проводится по той причине, что рабочих невозможно случайным образом распределить но двум фабрикам (только представьте себе расходы по перемещению рабочих, затраты на удовлетворение жалоб профсоюзов и т. д.). Независимая переменная принимает два значения: наличие или отсутствие свободного графика работы, а зависимой переменной является особым образом измеряемый показатель производительности труда. Предположим, что окончательный вариант плана следующий:

Фабрика в Питтсбурге:	Предварительное тестирование	Средняя производительность труда за 1 месяц до введения свободного графика работы
	Экспериментальное воздействие	Свободный график работы введен на 6 месяцев
	Заключительное тестирование	Средняя производительность труда в течение 6 месяцев работы по свободному графику
Фабрика в Кливленде:	Предварительное тестирование	Средняя производительность труда за 1 месяц до введения свободного графика работы в Питтсбурге
	Экспериментальное воздействие	Отсутствует
	Заключительное тестирование	Средняя производительность труда в течение 6 месяцев работы по свободному графику на фабрике в Питтсбурге

Результаты

На рис. 10.3 приведены четыре результата данного квазиэксперимента. Все графики показывают увеличение производительности труда на фабрике в Питтсбурге. Вопрос в том, вызвано ли это увеличение экспериментальной программой или другим (другими) фактором (факторами). Прежде чем продолжить чтение, попробуйте определить, какой из графиков лучше всего демонстрирует, что введение свободного графика работы увеличивает производительность труда, а также попытайтесь выявить факторы, угрожающие внутренней валидности исследования и затрудняющие интерпретацию результатов.

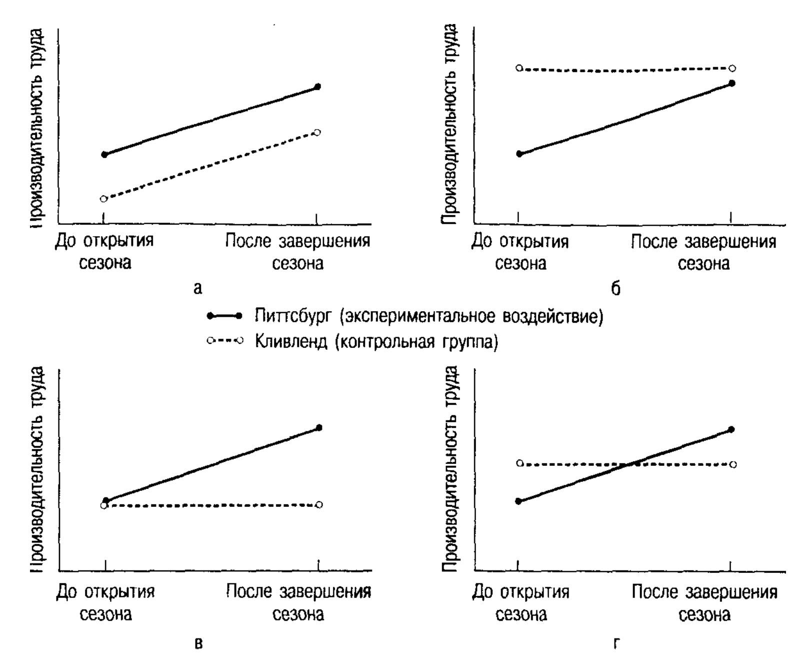


Рис. 10.3. Гипотетические результаты исследования по плану с неэквивалентными группами

Я думаю, что вам не составило труда увидеть, что на увеличение производительности труда кроме свободного графика работы повлияли какие-то другие факторы. Этот график говорит о важности использования контрольной группы, даже если она является неэквивалентной. Да, действительно, питтсбургская производительность труда увеличилась, но то же самое произошло и в Кливленде, а следовательно, данные о росте производительности в Питтсбурге могут быть связаны не с успехами экспериментальной программы, а с различными факторами, угрожающими внутренней валидности, которые вы изучали ранее. История и взросление — одни из таких возможных факторов. Может быть, в период между предварительным и заключительным тестированием проводились национальные выборы и рабочие разных предприятий испытали связанный с этим эмоциональный подъем, что привело к росту производительности труда. А возможно, производительность увеличилась с приобретением опыта.

Рисунок 10.3, б говорит о том, что производительность в Кливленде поддерживалась высокой на протяжении всего исследования, а в Питтсбурге сначала была очень низкой, но затем увеличилась вследствие введения свободного графика работы. Однако здесь есть две проблемы. Во-первых, результаты, полученные в Кливленде, могут отображать эффект потолка (см. главу 8). Это значит, что уровень производительности изначально был настолько высок, что эксперимент не смог показать его дальнейшее увеличение. Если бы можно было увидеть рост производительности (т. е. если бы значения на оси Y могли увеличиться), то вы бы увидели две параллельные линии, как на рис. 10.3, а. Вторая проблема заключается в том, что поскольку начальная производительность в Питтсбурге была настолько низкой, ее увеличение вызвано скорее эффектом регрессии, чем экспериментальной программой.

Рисунок 10.3, в на первый взгляд кажется отображением идеальных результатов. Обе группы начинают с одного уровня производительности труда, но ее рост наблюдается только в экспериментальной группе (Питтсбург). Такой результат может действительно иметь место, и исследователи будут просто счастливы его получить, но тем не менее возможна проблема. Поскольку две группы неэквивалентны, эффект отбора субъектов может сочетаться с влиянием других факторов. Это значит, что определенный фактор, например история или взросление может воздействовать на рабочих одной фабрики и не влиять на рабочих другой. Например, несложно представить здесь проблему отбор χ история: некоторое событие повлияло на рабочих Питтсбурга, а на рабочих Кливленда не оказало никакого влияния. Возможно, на питтсбургских рабочих повлияло то, что они знали об участии в исследовании (помните Хоторн?). А может быть, в период между предварительным и заключительным тестированием Стилеры выиграли кубок *Super Bowl* по американскому футболу, а так как рабочие в Питтсбурге очень активно болеют за свою команду, то общее ощущение успеха могло повысить у них рабочий настрой, а следовательно, и производительность труда. А Брауны, которые никогда не получали кубок, не смогут вызвать рост продуктивности у рабочих Кливленда.

Вы, вероятно, заметили сходство рис. 10.3, в и 10.1, иллюстрирующих исследование метода когнитивных интервью. Фишер и др. (Fisher et al., 1989) не проводили случайного распределения полицейских по группам, одной, в которой проводилось обучение по данному методу, и второй, контрольной, но исследователи считали обе группы приблизительно эквивалентными, так как применяли процедуру уравнивания с использованием «информации, собранной в... ходе предварительных опросов и сведений, полученных от вышестоящих офицеров полиции» (Fisher et al. 1989, p. 724). Таким образом, они считали, что результаты, приведенные на рис. 10.1, отображают реальное влияние обучения и что эффект отбора был минимальным.

Результаты, представленные на рис. 10.3, г, лучше всего подтверждают эффективность программы. В этом случае производительность труда у группы воздействия в начале исследования *ниже,* чем у контрольной группы, а в конце становится *выше,* чем у нее. Регрессию как причину улучшения можно отбросить, так как она возможна только до уровня контрольной группы, но не выше. Конечно, нельзя совсем исключить проблему отбора и возможность взаимодействия между отбором и другими факторами, но такой вариант эффекта взаимодействия в целом расценивается как подтверждение эффективности программы (Cook & Campbell, 1979).

Регрессия и уравнивание

При попытке снизить неэквивалентность групп с помощью процедуры уравнивания появляется особый вид угрозы внутренней валидности исследования. В только что рассмотренном примере эта проблема не возникла, но в прикладных исследованиях она иногда появляется. Уравнивание описывалось в главе 6 как альтернатива случайному распределению, подходящая для создания эквивалентных групп в случае управляемой независимой переменной. При использовании уравнивания участников разделяют на пары в зависимости от значения переменной уравнивания, а затем случайным образом распределяют по группам (см. о проце-

дуре уравнивания в главе 6). Однако, если при использовании плана с неэквивалентной контрольной группой две группы были набраны из популяций, имеющих значимые отличия друг от друга в отношении переменной уравнивания, уравнивание может вызвать проблемы. Использование процедуры уравнивания в этом случае может усилить **эффект регрессии** и даже привести к отказу от эффективной программы. Рассмотрим гипотетический пример.

Предположим, вы разрабатываете программу развития навыков чтения у плохо читающих детей. Вы объявляете о наборе добровольцев для участия в новой обучающей программе и отбираете тех, кто наиболее нуждается в помощи (т. е. детей с очень низкими средними результатами). Чтобы держать под контролем социоэкономпческий класс, вы набираете добровольцев, живущих в таких же районах, но в других городах, и создаете из них контрольную группу. Главная задача заключается в том, чтобы создать группы с одинаковыми начальными навыками чтения, поэтому вы решаете уравнять группы в отношении этой переменной. Вы проводите предварительный тест на навыки чтения у добровольцев из изучаемого района и у потенциальных членов контрольной группы, а затем по полученным результатам формируете две группы с одинаковыми средними оценками. Допустим, что результаты теста варьируются от 0 до 50 баллов. Вы решаете для обеих групп отобрать детей со средней оценкой 25 баллов. Далее группа воздействия занимается но разработанной программе, а контрольная группа не подвергается воздействию. Используется обычный план с неэквивалентной группой:

Экспериментальная группа	Предварительное тестирование	Программа ио чтению	Заключительное тестирование
Контрольная группа	Предварительное тестирование	—	Заключительное тестирование

Вы уверены, что ваша программа совершенна и действительно способна помочь множеству детей, и с нетерпением ждете результатов исследования. Поэтому вы шокированы, когда получаете следующие данные:

Экспериментальная группа	Предварительное тестирование = 25	Программа по чтению	Заключительное тестирование - 25
Контрольная группа	Предварительное тестирование = 25	-	Заключительное тестирование = 29

По-видимому, разработанная программа не просто не работает, но даже препятствует развитию навыков чтения — у контрольной группы обнаружено заметное улучшение! Что же произошло?

Очень вероятно, что эффект регрессии, вызванный процедурой уравнивания, перекрыл эффект экспериментального воздействия. Не забывайте, что экспериментальная группа была сформирована из детей со слабыми навыками чтения как наиболее нуждающихся в этой программе. Если провести тестирование всех детей, попадающих в эту категорию (т. е. в эту популяцию), то средняя оценка может оказаться довольно низкой, скажем 17. Однако при использовании процедуры уравнивания вы были вынуждены отобрать детей с гораздо более высокими оценками, чем в среднем у детей из популяции «плохо читающих». Возможно, что некоторые

дети из этой группы в ходе предварительного тестирования получили оценки выше, чем они должны были бы получить, ведь никакой тест не обладает абсолютной надежностью и могла возникнуть некоторая ошибка измерений. Поэтому при заключительном тестировании многие из этих детей получили более низкие оценки вследствие регрессии к среднему. Предположим, что программа была действительно эффективной и увеличивала оценку навыка чтения в среднем на 4 балла. Однако если эффект регрессии состоял в снижении оценки в среднем на 4 балла, то в результате будет обнаружено отсутствие изменений результатов от предварительного к заключительному тестированию:

$$25 + (+4) + (-4) = 25.$$

С участниками контрольной группы могло произойти обратное. Возможно, средняя оценка для их популяции была намного выше, чем 25 (например, 35) и они довольно хорошо читали (т. е. образовывали популяцию, отличающуюся от экспериментальной группы). Отбор участников, чьи оценки гораздо ниже, чем в среднем у популяции, проведенный с той целью, чтобы результаты предварительного тестирования соответствовали оценкам экспериментальной группы, мог привести к тому, что из-за эффекта регрессии оценки заключительного тестирования оказались выше. На результаты этих детей, полученные при заключительном тестировании, повлиял такой же эффект регрессии, заключающийся в изменении оценок на 4 балла. Таким образом:

$$25 + 0 + (+4) = 29.$$

На рис. 10.4 эта проблема представлена наглядно. Регрессия и вызванные программой улучшения взаимно исключили друг друга в экспериментальной группе, а на контрольную группу повлияла только регрессия, сдвинувшая оценки в сторону увеличения. В целом, программа по чтению могла быть весьма хорошей, но процедура уравнивания вызвала эффект регрессии, сделавший незаметной ее эффективность[1].

Этот вид регрессии обнаружил свой эффект при первой крупномасштабной попытке оценить эффективность программы «Рывок» (60-х гг. XX в.) — одной из центральных программ законодательной инициативы президента Линдона Джонсона «Великое общество» (Campbell & Erlebacher, 1970). Программа была начата в 1965 г. и представляла собой масштабную попытку дать возможность дошкольникам из бедных семей сделать «рывок» к школе: развить у детей различные, связанные со школой навыки, а также привлечь к этому процессу их родителей. К 1990 г. около 11 млн детей участвовали в программе, и сегодня «Рывок» считается самой успешной социальной программой, проводимой правительством (Horn, 1990). В начале 70-х гг. ее критиковали за неспособность вызывать продолжительный эффект. Критики преимущественно основывались на «Вестингхаузовском исследовании» (названном так, поскольку его позволил начать грант, выданный Вестингхаузовской корпорации обучения и Университету штата Огайо), проведенном Виктором Цицирелли и его коллегами (Cicirelli, V.G., Cooper, W.H. & Granger, R.L., 1969).

[1] Хотя проведение прикладного исследования в нолевых условиях может предотвратить подобную проблему, лучше всего проверить навыки чтения у большой группы детей, уравнять их по результатам данного теста, а затем случайным образом распределить их на экспериментальную и контрольную группы.

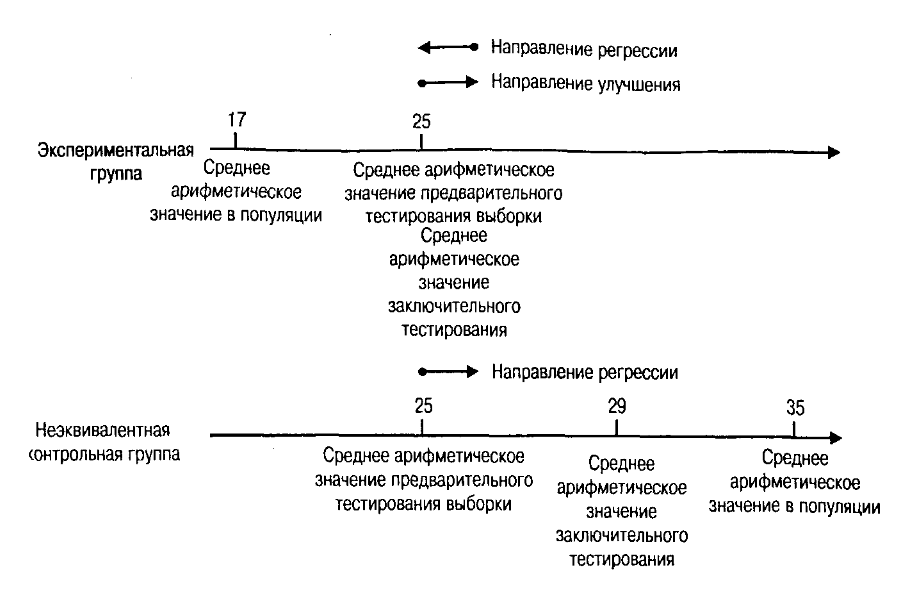


Рис. 10.4. Гипотетическое влияние эффекта регрессии при уравнивании неэквивалентных групп

В ходе Вестпнгхаузовского исследования был обнаружен «эффект затухания» — раннее развитие способностей у детей, произошедшее под влиянием программы «Рывок», к третьему классу стиралось. Вывод, конечно, был сделан такой, что федеральные доллары тратятся на неэффективную социальную программу. Именно на это в своем обращении к конгрессу указал президент Никсон, ссылаясь на вестингхаузовское исследование. Впоследствии в годы правления Никсона финансирование «Рывка» неоднократно ставилось под сомнение. В то же самое время главный фундамент критики, вестннгхаузовское исследование, получил отрицательную оценку социологов.

Из-за того, что программа «Рывок» уже давно действовала к моменту, когда стартовал вестингхаузовский проект по ее оценке, в ходе его было невозможно случайным образом распределить детей на экспериментальную и контрольную группы. Вместо этого исследовательская группа набрала группу детей, участвующих в программе «Рывок», и уравняла их по когнитивным показателям с группой детей, не принимающих участия в этой программе. Но для того чтобы уравнять детей по когнитивным показателям, из участников программы «Рывок» были отобраны дети, чьи оценки заметно превышали среднее значение, и контрольные дети, чьи оценки были заметно ниже средних для их группы. А это в точности повторяет ситуацию, только что описанную в приведенном гипотетическом примере программы развития навыков чтения. Таким образом, отсутствие улучшений в третьем классе у группы из программы «Рывок» было, но крайней мере отчасти, связано с регрессией, вызванной процедурой уравнивания (Campbell & Erlebacher, 1970).

В защиту вестингхаузовской исследовательской группы следует отметить, что они отчаянно возражали политикам, которые собирались закрыть программу. Цицирелли (Cicirelli V.G., 1984) настаивал на том, что исследование *«не показало, что "Рывок" был неудачным»* (р. 915; курсив автора), что необходимы дополнительные

исследования, а также что должны быть предприняты «серьезные усилия по расширению и усовершенствованию этой программы» (р. 916). Не так давно (Cicirelli, V.G., 1993) Цицирелли отметил, что исходя из вестингхаузовского исследования можно порекомендовать «не останавливать программу "Рывок", а на основании открытии, сделанных в ходе применения этой программы, постараться сделать ее более эффективной» (р. 32).

Планы с неэквивалентной контрольной группой не всегда приводят к разногласиям, подобным тем, которые возникли в ходе вестингхаузовского исследования. Нижеследующий пример демонстрирует попытку поднять самооценку бейсболистов малой лиги с помощью особого курса занятий для тренеров.

Пример 24. План с неэквивалентной контрольной группой

Тренер может играть очень важную роль в жизни детей, занимающихся спортом. На детей, у которых только начали формироваться навыки, тренер может оказать весьма деструктивное влияние, особенно если единственная его цель победа в чемпионате, а любые недостатки в игре детей вызывают его порицание и насмешки. Но тренеры могут быть и прекрасными учителями, а также могут служить примером для детей в годы их становления. Смолл, Смит, Барнетт и Эверетт (Smoll, Smith, Barnett & Everett, 1993) провели исследование по оценке программы подготовки тренеров, призванной помочь им в формировании самоуважения у своих юных спортсменов. За две недели до открытия бейсбольного сезона в малой лиге несколько тренеров прошли программу, названную «Тренинг эффективности для тренеров». В программе были выделены несколько целевых способов поведения, которых необходимо придерживаться тренерам: поощрение усилий наравне с хорошей игрой, ободрение и инструкции о правильных действиях, если ребенок совершил ошибку, поощрение действий, направленных на коллективную игру команды. Программа оценивалась в ходе исследования, проведенного по плану с неэквивалентной контрольной группой. В экспериментальную группу вошли восемь тренеров из одной лиги, а в контрольную — десять тренеров из разных лиг. Почему нельзя было использовать экспериментальный план, набрать тренеров из одной лиги и случайно распределить их по группам? На это есть три причины. Во-первых, при использовании плана с неэквивалентной контрольной группой легче координировать тренинговую программу. Во-вторых, такое решение позволило, чтобы количество побед и поражений в каждой группе составило 50%. А в-третьих, это обеспечило «отсутствие утечки информации и передачи сведений о программе тренерам в контрольной группе» (р. 603). Таким образом, две группы были неэквивалентны по определению, но исследователи постарались сделать их как можно более похожими, за исключением использования тренинга. Например, они контролировали социоэкономическое положение тренеров.

Дети, занимающиеся у тренеров из экспериментальной и контрольной групп, были исследованы до и после сезона малой лиги. Мальчики прошли стандартизованный тест на самооценку, а в ходе заключительного тестирования было спрошено их мнение о действиях тренеров (исследователи, проводящие интервью, не знали, к какой из групп принадлежали мальчики, к экспериментальной или контрольной). Оказалось, что игрокам из экспериментальной группы тренеры и занятия

бейсболом нравятся больше, чем детям из контрольной группы. В отношении самооценки различий между двумя группами обнаружено не было, но когда Смолл и др. (Small et al., 1993) исследовали только мальчиков, чьи оценки, полученные при заключительном тестировании, лежали ниже медианы, они получили зависимость, представленную на рис. 10.5. Хорошо видно, что если с мальчиками с относительно низкой самооценкой занимался тренер из экспериментальной группы, то в начале сезона у них был обнаружен заметный рост этого показателя, а если тренер был из контрольной группы, то самооценка не менялась (видимое снижение не было значимым). Если допустить, что мальчики с низкой самооценкой легче других падают духом при спортивных неудачах, то такие результаты покажутся весьма впечатляющими. Если тренер правильно обращается с детьми, он может оказать положительное влияние на жизнь маленьких игроков. Это действительно хорошее исследование.

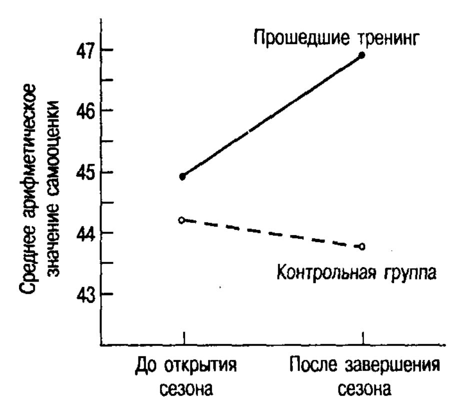


Рис. 10.5. Изменение самооценки у игроков малой лиги до и после тогс, как тренеры были обучены способам повышения этой характеристики (из работы Small et al., 1993)

В планах с неэквивалентной контрольной группой обычно используется предварительное тестирование, но иногда обстоятельства складываются так, что его провести невозможно. Одним из таких непредвиденных обстоятельств явилось землетрясение, произошедшее в Сан-Франциско в 1989 г. Джеймса Вуда и Ричарда Буцина из Университета штата Аризона это событие навело на мысль об проведении исследования ночных кошмаров — тема, давно интересовавшая обоих исследователей (Wood & Bootzin, 1990). Вместе со своими коллегами из Стэнфордского университета (расположенного неподалеку от эпицентра землетрясения) они быстро разработали исследование, чтобы выяснить, влияет ли переживание такого травмирующего события на содержание снов в целом и ночные кошмары в частности (Wood, Bootzin, Rosenhan, Nolen-Hoeksema & Jourden, 1992). Было необходимо использовать план с неэквивалентной контрольной группой. Как это обычно происходит в исследованиях с таким планом, группы были неэквивалентны с самого начала (студенты из двух различных штатов), кроме того, на одну группу было оказано воздействие одного типа (они пережили землетрясение), а на другую — другого (отсутствие непосредственного переживание землетрясения).

Экспериментальная группа состояла из студентов Стэифорда и Государственного университета города Сан-Хосе, переживших землетрясение. Неэквивалентная контрольная группа была сформирована из студентов Университета штата Аризона. Они не столкнулись с землетрясением, но находились под влиянием многочисленных сообщений средств массовой информации. Все участники записывали содержание своих снов, которые затем анализировались по тематическому содержанию и частоте появления кошмаров. Были получены очень интересные результаты. Через три недели после начала исследования по меньшей мере один кошмар видели около 40% участников, перенесших землетрясение, и только 5% контрольных испытуемых. Из всех ночных кошмаров, пережитых первой группой, приблизительно одна четверть была о землетрясениях. Практически никто из контрольной группы не видел таких кошмаров. Кроме того, частота кошмаров значимо коррелировала с уровнем тревоги, испытанной участниками в момент землетрясения.

Прекрасно зная о проблемах интерпретации квазиэкспериментальных исследований, Вуд и др. (Wood et al., 1992) понимали трудность сравнения неэквивалентных групп. Например, не имея полученной при предварительном тестировании (до землетрясения) информации о частоте ночных кошмаров, они не смогли «исключить возможность того, что жители Калифорнии видят по ночам больше кошмаров о землетрясениях, чем жители Аризоны, даже если в ближайшее время не было зафиксировано ни одного подобного явления» (р. 222). Возможно, что кошмары о землетрясениях — это обычное дело для жителей Калифорнии. Однако, полагаясь отчасти на свой исследовательский опыт в области ночных кошмаров, авторы пришли к выводу, что частота кошмаров была особенно высока в группе из Калифорнии и, по всей видимости, была связана с их недавними травматическими переживаниями.

Планы с прерванной временной последовательностью

Если бы Вуд и его коллеги могли предвидеть землетрясение в Сан-Франциско, они бы начали собирать данные о ночных кошмарах у своих участников за несколько месяцев до этого события и продолжали бы делать это еще несколько месяцев после него. Это помогло бы им определить: а) действительно ли землетрясение увеличивает количество ночных кошмаров у людей, живущих в зоне землетрясения, и б) действительно ли частота кошмаров достигает пика сразу после землетрясения, а затем возвращается к обычному уровню. Конечно, даже талантливые психологи-исследователи не могут предсказать землетрясение, а поэтому Вуд и его сотрудники сделали все, что могли, и разработали исследование с неэквивалентной контрольной группой. Если бы они могли проводить измерения какое-то время до и после воздействующего на поведение события, то план их исследования носил бы название **плана с прерванной временной последовательностью.**

В соответствии с обозначениями, использованными Кэмпбеллом и Стэнли (Campbell & Stanely, 1963), типичное исследование с прерванной временной последовательностью можно представить следующим образом:

$$И_1\ Н_2\ Н_3\ Н_4\ Н_5\ В\ Н_6\ Н_7\ Н_8\ Н_9\ Н_{10},$$

где все Н означают измерения, проведенные до и после В — момента, в который была начата программа воздействия или произошло определенное событие (напри-

мер, землетрясение). В — это «прерывание» в прерванной временной последовательности. Конечно, количество измерений, проводимых до и после В, варьируются от исследования к исследованию и их число не ограничено пятью для каждого из этапов. Также не обязательно, чтобы количество измерений до и после прерывания было одинаковым. Как правило, чем больше данных собрано в различные моменты, тем лучше, и некоторые специалисты (например, Orwin, 1997) советуют проводить по крайней мере 50 измерении до момента прерывания.

Результаты

Основное преимущество плана с прерванной временной последовательностью заключается в том, что он позволяет оценивать тенденции — предсказуемые события, которые возникают с течением времени. Предположим, к примеру, что вы хотите обнаружить влияние двухмесячной антиникотиновой кампании на количество курящих подростков. Программа может включать определенные методы убеждения, личные беседы, показ подросткам изображений легких курильщиков и т. д. Допустим, что у вас есть хороший способ выявления курильщиков. Вы проводите измерения за месяц до и через месяц после антиникотиновой программы и получаете результаты, представленные на рис. 10.6.

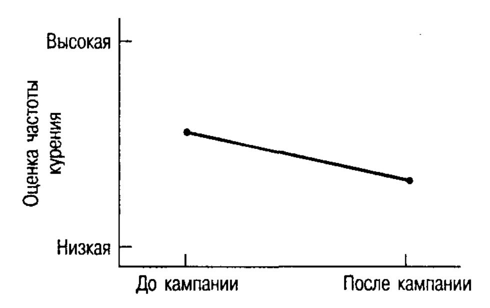


Рис. 10.6. Частота курения непосредственно до и после гипотетической антиникотиновой кампании

Успешна ли программа? Снижение частоты курения от предварительного тестирования к заключительному очевидно, но оценить его при отсутствии контрольной группы довольно сложно. Но даже без контрольной группы можно понять, эффективна ли программа, если провести не одно, а несколько измерений до и после ее проведения. На рис. 10.7 представлены возможные результаты исследования по оценке антиникотиновой кампании, в ходе которого частота курения измерялась каждый месяц в течение года до и года после проведения программы.

На рис. 10.7, а ясно показано, как с помощью плана с прерванной временной последовательностью можно выявить тенденцию. В данном случае видно, что уменьшение числа курильщиков, которое выглядело так замечательно на рис. 10.6, — это не более чем общая тенденция к снижению числа курильщиков среди взрослых. Это говорит о важной особенности планов с прерванной временной последовательностью — они позволяют исключить (т. е. фальсифицировать, помните?) альтернативные объяснения изменений, произошедших от предварительного к заключительному тестированию.

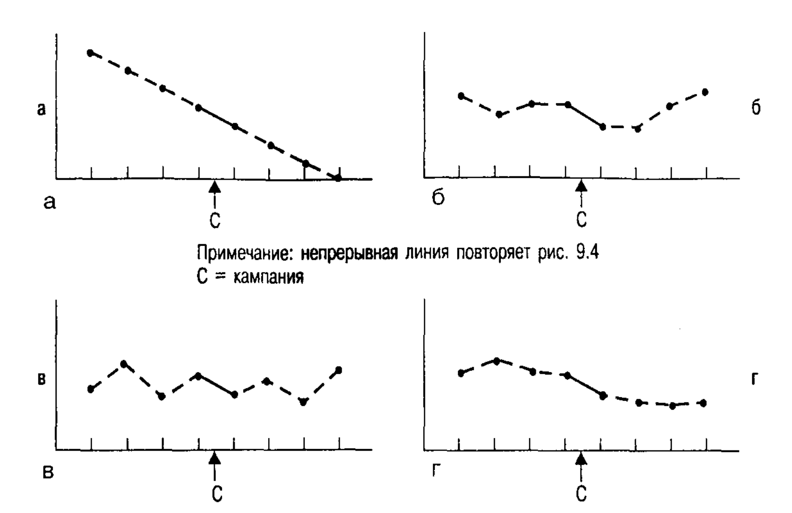


Рис. 10.7. Гипотетическая антиникотиновая кампания с планом с прерванной временной последовательностью - несколько возможных результатов

На рис. 10.7, б представлены два других результата, вызывающие вопросы об эффективности программы. На рис. 10.7, б количество курильщиков держалось довольно постоянным до кампании, а затем снизилось, но лишь немного. Другими словами, если антиникотиновая программа вообще имела какой-нибудь эффект, он был краткосрочным. На рис. 10.7, в снижение числа курильщиков после проведения программы является частью другой общей тенденции, па этот раз — периодических колебаний между более высоким и более низким уровнем курения. Идеальный результат показан на рис. 10.7, г. Здесь число курильщиков постоянно и довольно велико до начала программы, падает после того, как антиникотиновая программа приводится в исполнение, и сохраняется низким некоторое время после ее окончания. Обратите внимание, что показанный на рис. 10.7, г относительно неизменный базовый уровень, выявленный до начала кампании, позволил исследователям исключить эффект регрессии.

Пример 25. План с прерванной временной последовательностью

Результат, аналогичный последнему, был получен в исследовании производительности труда, проведенном в железолитейном цехе Вагнером с соавторами (Wagner, Rubin & Callahan, 1988). Исследователи изучали влияние введения стимулирующего плана, в соответствии с которым рабочие рассматривались не индивидуально, а как члены небольших групп, каждая из которых несла ответственность за определенную линию производства. Данные о производительности труда собирались в течение 10 лет, 4 года до введения стимулирующего плана и 6 лет после этого. Как видно из графика на рис. 10.8, где по оси X отложены отрезки времени, производительность до введения плана была довольно стабильной и не производила большого впечатления, но после введения плана она стала равномерно увеличиваться.

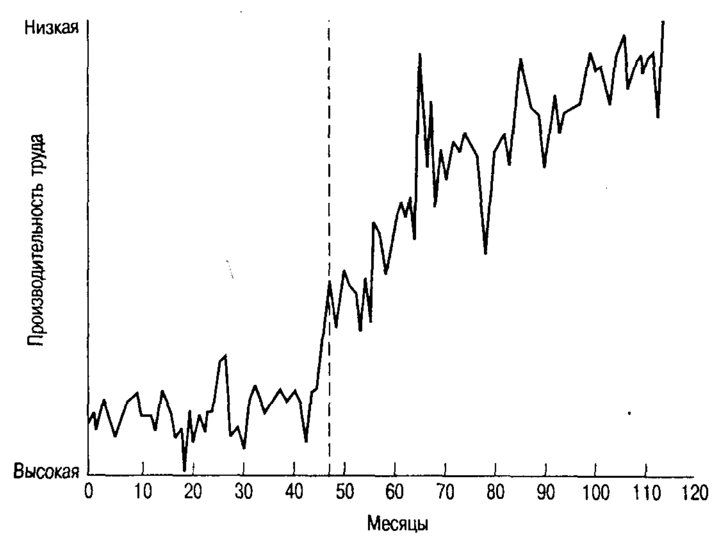


Рис. 10.8. План с прерванной временной последовательностью; влияние стимулирующего плана на производительность труда рабочих железолитейного цеха (из работы Wagner et al., 1988)

Это исследование также показывает, как ученые, использующие план с прерванной временной последовательностью, справляются с факторами, угрожающими внутренней валидное™ их работы. Рисунок 10.8 определенно показывает, что стимулирующий план сотворил чудо, но на обнаруженные изменения могли повлиять другие факторы, в том числе история, инструментарий и отбор. Авторы данного исследования указали, что история не повлияла на изменения, так как, тщательно изучив практически все события, имевшие место в период до и после произошедших изменений, они не обнаружили повода подозревать, что некоторое необычное событие привело к скачку производительности труда. В действительности определенные события, от которых можно было ожидать негативного влияния на производительность труда (например, спад в автомобильной индустрии, повлиявший на продажи отливаемых из железа изделий), не оказали подобного влияния. Также исследователи исключили влияние инструментария: он мог представлять проблему, если бы значения среднего арифметического производительности труда рабочих с годами изменялось. Но этого не произошло. В-третьих, хотя мы обычно думаем, что отбор испытуемых является потенциальным осложнителем только в исследованиях с двумя и более неэквивалентными группами, эта проблема может проявиться и в плане с временной последовательностью, если во время действия нового плана будет обнаружена значимая текучесть кадров. В таком случае коллектив рабочих до внедрения нового плана будет отличаться от группы, работающей после этого события. В исследовании Вагнера и др. этого не произошло. В целом, планы, подобные использованному в данном исследовании, чувствительны к различным факторам, угрожающим внутренней валидности, но их можно исключить, если систематически изучать доступную информацию, как это делали Вагнер и его коллеги.

Разновидности планов с временной последовательностью

Иногда выводы, сделанные на основании исследовании, проведенных по плану, с прерванной временной последовательностью, можно укрепить, проведя особое контрольное сравнение. Один подход заключается в объединении самых полезных особенностей плана с неэквивалентной контрольной группой (контрольной группы) и плана с прерванной временной последовательностью (долговременного анализа тенденций). Такой план можно представить следующим образом:

$$\begin{array}{cccccccccccc} H_1 & H_2 & H_3 & H_4 & H_5 & \mathbf{B} & H_6 & H_7 & H_8 & H_9 & H_{10} \\ H_1 & H_2 & H_3 & H_4 & H_5 & & H_6 & H_7 & H_8 & H_9 & H_{10} \end{array}$$

Если забежать вперед и посмотреть на рис. 10.10 из вставки 10.2, можно увидеть классический пример применения данной стратегии: исследование по оценке мер, применяемых в штате Коннектикут для снижения скорости движения автомобилей, путем сравнения результатов, полученных в этом штате, с данными из других штатов.

Вторая стратегия по укреплению выводов исследования с временной последовательностью применяется, если программа вводится в разных местах в разное время. Такой план Кук и Кэмпбелл (Cook & Campbell, 1979) назвали **планом с прерванной временной последовательностью с переключающими повторениями** и обозначили следующим образом:

$$\begin{array}{cccccccccccc} H_1 & H_2 & H_3 & \mathbf{B} & H_4 & H_5 & H_6 & H_7 & H_8 & H_9 & H_{10} \\ H_1 & H_2 & H_3 & H_4 & H_5 & H_6 & H_7 & \mathbf{B} & H_8 & H_9 & H_{10} \end{array}$$

В соответствии с этой процедурой одно и то же воздействие или программа проводится в двух различных местах в разное время. Контрольная группа не используется, но этот план обладает всеми достоинствами планов с повторениями. Если результаты, полученные в месте 2, соответствуют результатам места 1, то исследователи имеют больше оснований для уверенности в универсальном характере изучаемого явления. Такой результат был получен в неопубликованном исследовании, отчет о котором представлен у Кука и Кэмпбелла (Cook & Campbell, 1979). Это исследование было проведено в конце 40-х — начале 50-х гг. XX в., когда телевидение только начинало оказывать влияние на нашу жизнь. Нескольким сообществам штата Иллинойс было выдано разрешение на постройку телевизионных станций, но с 1951 по 1953 г. разрешения не выдавались. Это позволило исследователям изучить влияние появления телевидения на сообщества в два различных момента времени: в конце 40-х гг., до прекращения выдачи разрешений, и сразу после 1953 г., когда их снова ввели в употребление. Предполагая, что новое изобретение снизит процент читающих, исследователи изучили данные о потоке людей в библиотеках и получили подтверждение своей теории. С момента проникновения телевидения в жизнь поток людей в библиотеках снизился, причем оба изучаемые момента времени практически не отличались в этом отношении.

Третий подход к проведению исследований по плану с прерванной временной последовательностью без использования контрольной группы заключается в измерении нескольких зависимых переменных, причем предполагается, что одни из них находятся под влиянием «прерывания», а другие остаются без изменений. Имен-

но эта стратегия была использована в исследовании Штольценберга и Далессио (Stolzenberg & D'Alessio, 1997). Ученые решили изучить влияние калифорнийского закона об окончательном приговоре к тюремному заключению — так называемой политики «три удара, и ты вылетел», — на рост преступности. Суть этой политики состоит в том, что человек, признанный виновным в трех серьезных уголовных преступлениях, автоматически наказывается тюремным заключением. Собрав данные по 10 крупнейшим городам штата Калифорния, Штольценберг и Далессио исследовали два вида преступности (т. е. две зависимые переменные). Они изучили серьезные преступления, число которых предположительно должно было снизиться благодаря введению закона об окончательном приговоре, и относительно несерьезные преступления (проступки), на которые новый закон не должен был повлиять. На рис. 10.9 показаны полученные результаты, которые отражают достоинства использования временных последовательностей. Если посмотреть на кривую серьезных преступлений после введения закона, то обнаружится снижение уровня преступности, особенно в сравнении с горизонтальной линией несерьезных преступлений. Но если посмотреть на кривую серьезных преступлений в целом, то станет ясно, что снижение числа серьезных преступлений является частью тенденции, обозначившейся около 1992 г. В целом исследователи обнаружили, что закон «о трех ударах» не оказал заметного влияния ни на один из видов преступности.

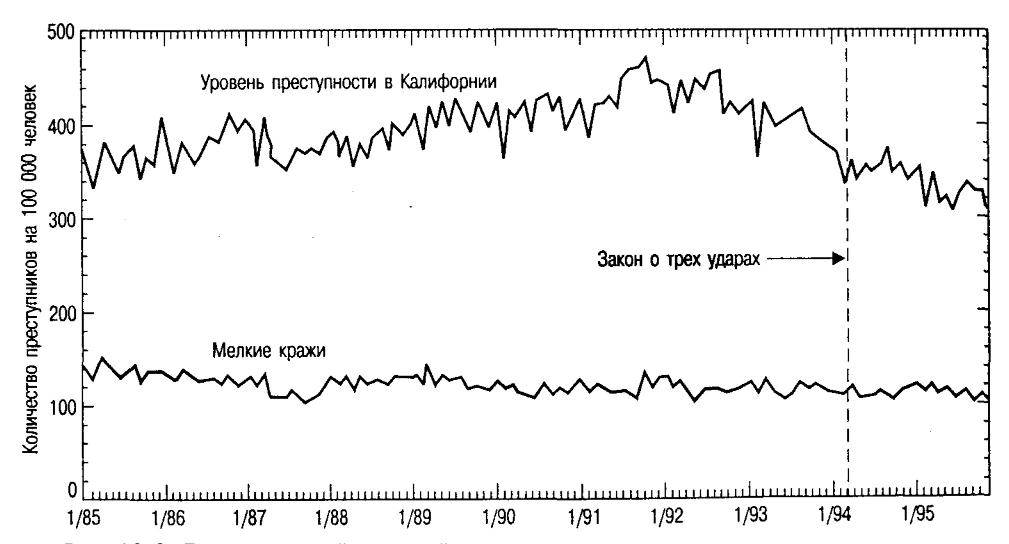


Рис. 10.9. План с прерванной временной последовательностью с использованием двух зависимых переменных; влияние закона об окончательном приговоре на уровень преступности (из работы Stolzenberg & D'Alessio, 1997)

Оценка программ

Попытка оценить изменения, вызванные законом о тюремном заключении, — это лишь один пример прикладного исследования, носящего название **оценки программы.** Эта область исследований стала разрабатываться в 60-х гг. в связи с необ-

ходимостью оценки таких социальных программ, как «Рывок». Данный вид исследований призван не просто ответить па вопрос «Эффективна ли программа X?», но имеет гораздо более широкие цели. В целом, оценка программы включает: а) процедуры для выявления необходимости в конкретной программе п определения того, кто получит пользу от ее введения, б) оценку того, проводится ли программа в соответствии с планом, и если нет, то какие изменения можно сделать, чтобы облегчить ее реализацию, в) методы для оценки результатов проведения программы и г) анализ затрат с целью определить, оправдывает ли полученная польза произведенные финансовые вложения. Рассмотрим более подробно каждый из этих пунктов. Но сначала вам необходимо прочитать вставку 10.2, посвященную документу, составленному Дональдом Кэмпбеллом в 1969 г. Этот документ считается одним из «самых важных документов в недолгой истории оценки программ».

ВСТАВКА 10.2

Классические исследования —реформы как эксперименты

Статья Дональда Кэмпбелла «Реформы как эксперименты», написанная им в 1969 г., примечательна по трем причинам. Во-первых, в ней убедительно доказывается, что нам необходимо выработать экспериментальный подход по отношению к социальным реформам. В самом начале статьи Кэмпбелл выступил в защиту применения научного образа мышления, написав, что «мы должны быть готовы к экспериментальной проверке социальных реформ, проверке, в ходе которой мы испытываем новые программы, направленные на решение особых социальных проблем, узнаем.эффективны ли эти программы, и на основании этого оставляем без изменений, копируем, модифицируем или отвергаем их» (р. 409).

Во-вторых, в статье Кэмпбелла было описано несколько исследований, ставших классическими в области оценки программ, и явившихся прототипами других исследований с использованием таких планов, как план с прерванной временной последовательностью. Вероятно, самым известным примером является его описание исследование по оценке мер, направленных на снижение скорости движения автомобилей в штате Коннектикут (Campbell & Ross, 1968). 1955 г. был годом рекордного количества смертей в автокатастрофах (324) и губернатор штата Коннектикут Абрахам Рибикофф ввел ограничения скорости, разумно предположив, что скорость движения и количество смертей в автокатастрофах коррелируют. В следующем году количество смертей упало до 284. Для Рибикоффа этого было достаточно, чтобы заявить, что на основании «сохранения 40 жизней в 1956 г. и снижения количества похоронных маршей... на 12,3% по сравнению с 1955 г. мы можем сказать, что эта программа действительно стоящая» (процитировано по Campbell, 1969, р. 412). Но действительно ли это так?

Я надеюсь, вы уже подумали о том, что здесь возможны и другие объяснения снижения уровня смертности. Например, мог сказаться фактор истории. Возможно, в 1956 г. погода была лучше, а дороги суше. Еще более вероятно влияние регрессии: 324 - это хороший пример максимальной оценки, вслед за которой обычно проходит регрессия к среднему. И действительно, Кэмпбелл указал, что на полученных в Коннектинуте результатах сказалась регрессия и что «в литературе, посвященной экспериментальным социальным реформам, регрессия - это, вероятно, наиболее часто встречающийся вид самомистификации» (р. 414). Это характерно для данного вида исследований, так как меры, подобные снижению допустимого предела скоро-

сти, часто принимаются сразу после того, как происходят особо трагические события. Теория вероятности говорит нам о том, что в следующем году дела не будут обстоять так же плохо. Только ли регрессия повлияла на полученные результаты? Вероятно, нет. Проведя исследование по плану с прерванной временной последовательностью и неэквивалентной контрольной группой (близлежащие штаты, в которых не было принято мер по снижению скорости автомобилей), Кэмпбелл сделал вывод, что эти меры имели некоторый эффект, хотя и не такой значительный, как думал губернатор. Результаты этого исследования представлены на рис. 10.10.

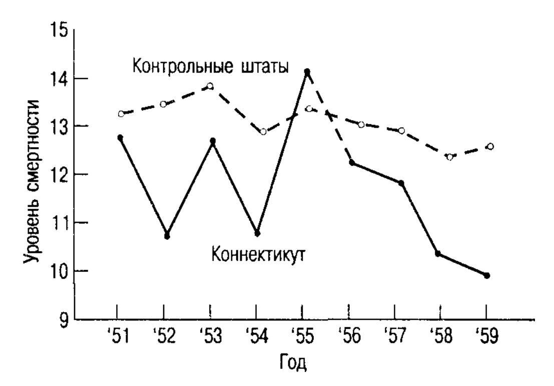


Рис. 10.10. Введение ограничения скорости в Коннектикуте - классический пример плана с прерванной временной последовательностью и неэквивалентной контрольной группой (из Campbell, 1969)

Третья причина особой важности статьи Кэмпбелла заключается в том, что она дает представление о политической реальности исследований, посвященных социальным вопросам. Политики часто предлагают программы, в эффективность которых они просто верят. Они говорят о необходимости тщательной оценки таких программ, но не слишком обрадуются, узнав о получении отрицательных результатов. Ведь поддержав программу, они оказываются заинтересованы в ее успехе и продолжении, особенно если она приносит пользу родному штату или округу этого политика. По этой причине политики и администрация, принятые на работу для проведения конкретной программы, редко настаивают на скрупулезной оценке и склонны довольствоваться положительными результатами, даже если они получены в ходе плохо разработанных исследований. Например, губернатору Рибикоффу было достаточно информации о смертности в результате автокатастроф, полученной непосредственно до и после принятия мер по ограничению скорости.

Кэмпбелл рекомендует изменить отношение к этому вопросу и сместить акцент с важности конкретной программы на важность проблемы. Это заставит политиков и администрацию думать о программах как о попытках экспериментального решения проблем. В таком случае, прежде чем выбрать одну эффективную, будут испытаны несколько различных программ. Как выразился Кэмпбелл в заключительной части своей статьи: «*Директивная администрация* настолько заранее уверена в эффективности реформ, что просто не может позволить себе честную оценку... *Экспериментирующая администрация* подтверждает эффективность реформ исходя из важности проблемы, а не на основании уверенности в своей правоте и готова рассмотреть другие возможные решения этой проблемы, если первое оказывается неудовлетворительным. Поэтому она не боится строгого анализа проводимых реформ») 1969, р. 428; курсив автора).

Планирование программ — анализ потребности

Учреждение начинает использовать определенную программу, так как его руководители находят, что в этом есть потребность. Как определяется эта потребность? Очевидно, что мнения руководства об осмысленности применения программы не достаточно. Программа физических упражнений для пенсионеров кажется весьма подходящей, но если никто не захочет в ней участвовать, то время и деньги будут потрачены зря. Прежде чем планировать любой проект, необходимо провести анализ потребности.

Анализ потребности — это набор процедур, с помощью которых определяется, существует ли достаточно большая популяция, которой будет полезна предложенная программа, может ли данная программа решить строго определенную проблему и будут ли члены этой популяции пользоваться программой. Существуют различные методы оценки необходимости, и очень важно прибегнуть по крайней мере к нескольким из них, поскольку необходимость очень легко переоценить. Один из поводов для осторожности — **переоценка по** запоминаемости, с которой вы познакомились в главе 1 при обсуждении типов познания. События, попадающие на первые страницы, привлекают наше внимание и лучше запоминаются. Поскольку они так легко закрепляются в памяти, мы склонны переоценивать частоту их появления. Достаточно было широкой огласки пары историй о детях, оставленных взявшими отпуск родителями, чтобы раздался звонок с предложением о проведении новой программы для решения этой весьма распространенной проблемы. Также потребность в новой программе может переоцениваться людьми, которым выгодно ее существование (она сохраняет им рабочие места).

Как указали Позавак и Кэри (Posavac & Carey, 1997), есть несколько способов определить потенциальную необходимость в программе.

- *Данные переписи населения.* Если предлагаемая вами программа предназначена для пожилых людей, а в данном сообществе живут лишь несколько пенсионеров, то очевидно, что программа будет мало полезной. Данные переписи населения (www.census.gov) дадут вам базовую демографическую информацию о количестве людей, соответствующих различным категориям. Кроме того, эта информация достаточно подробна, чтобы вы могли определить такие показатели, как количество матерей-одиночек в возрасте до 21 года, количество людей с различными физическими или психическими нарушениями или число пожилых людей, находящихся за чертой бедности. Кроме данных переписи местные организации могут иметь определенные архивные данные. Например, руководство корпорации, которая планирует организовать детский сад при производстве, может оценить количество сотрудников, имеющих маленьких детей. Также важную информацию о необходимости проведения программы могут дать опубликованные отчеты об исследованиях. Если исследование показало устойчивое снижение уровня интеллекта детей, живущих в нищете, то это говорит о необходимости программы по улучшению экономических условий жизни.

- *Изучение доступных услуг.* Нет необходимости начинать программу доставки пищи на дом инвалидам, если такая программа уже успешно действует. Таким образом, один из очевидных шагов в процессе анализа необходимости — это составление перечня существующих услуг, включая информацию о том, кто предоставляет услуги, какие именно услуги предоставляются, и оценку количества людей, которым предоставляются данные услуги. К примеру, вы можете обнаружить, что действующая программа доставки пищи проводится организацией, имеющей слишком большое финансирование и небольшой штат сотрудников, а это приводит к тому, что охвачено не более 10% нуждающихся. В таком случае разработка второй программы может оказаться весьма полезной.
- *Изучение потенциальных пользователей.* Третья стратегия анализа необходимости заключается в проведении опроса общественности: либо высоко репрезентативной выборки, либо целевой группы, определяемой по данным переписи населения. В ходе опроса можно выяснить мнение участников о необходимости проведения дайной программы. Но более важно, что опрос может помочь оцепить количество людей, которые действительно будут участвовать в предлагаемой программе.
- *Ключевые информаторы, фокус-группы и общие собрания.* **Ключевой информатор** — это человек, имеющий опыт и специальные знания о существующей проблеме, которая иначе не станет известна исследователю (Gilchrist & Williams, 1999). Это могут быть общественные активисты, священники, люди, входящие в правление различных социальных служб, и т. д. Иногда они могут стоять слишком близко к проблеме, чтобы оценивать ее реалистично, но в целом они представляют собой очень важный источник информации. **Фокус-группа** — небольшая (обычно 7-9 человек) группа людей, отвечающих на ряд вопросов, посвященных определенной теме, например необходимости в конкретной программе. Изучение мнения фокус-группы часто проводится после опроса общественности; предварительно это делают, чтобы уточнить вопросы, используемые при опросе общественного мнения. Члены фокус-групп обычно сходны по возрасту и социоэкономическому статусу, что делает группу в целом подходящей для выполнения определенного задания. Кроме того, полезную информацию иногда можно получить па **общих собраниях** — открытых встречах, куда приглашаются все члены сообщества, в котором хотят ввести определенную программу ключевые информаторы. Фокус-группы и общие собрания могут очень помочь в сборе информации, но исследователям необходимо с осторожностью подойти к этому занятию и не переоценить слова речистых информаторов, членов фокус-групп и лидеров общих собраний.

В последние два десятилетия во многих корпорациях распространилось мнение о том, что прибыль имеет связь со здоровьем рабочих. Поэтому компании принялись разрабатывать, реализовывать и проводить оценку программ по охране здоровья рабочих. В нижеследующем примере описывается крупномасштабная программа, которая началась с тщательного анализа потребности.

Пример 26. Оценка потребности

В компании «Дюпон» перед началом проведения программы, призванной улучшить здоровье рабочих, был проведен анализ потребности (Bertera, 1990). Программа предусматривала введение ряда изменений, касающихся 110 000 сотрудников, работающих на 100 заводах фирмы. Было ясно, что затраты на проведение такой масштабной программы будут огромны, а поэтому было важно четко продемонстрировать потребность в ней.

Проведенная компанией «Дюпон» оценка потребности включала анализ существующей информации о частоте различных видов заболеваний сотрудников фирмы, о причинах смерти рабочих, а также об отсутствии на рабочих местах и нетрудоспособности сотрудников за последние 15 лет. Был получен один непредвиденный результат: обнаружилось, что сотрудники, выполняющие самую низкооплачиваемую работу, имели самые высокие показатели по всем основным категориям заболеваний. Это открытие показало исследователям, что данная группа рабочих требует особого внимания.

Еще одно свидетельство необходимости программы по охране здоровья было получено при изучении действующих в фирме программ по улучшению здоровья сотрудников. Исследование выявило широкий набор программ, проводимых медицинским персоналом заводов фирмы «Дюпон», в том числе программы по снижению веса, антиникотиновые и антистрессовые программы и др. Однако в основном они заключались в проведении лекций и показе фильмов без последующей работы с сотрудниками и без систематической оценки эффективности этих программ. Также был проведен опрос сотрудников, чтобы выявить их знания об охране здоровья, оценку своей деятельности по оздоровлению, готовность изменить привычки питания и предпочтения по отношению к различным видам программ по охране здоровья.

На основании полученной информации в фирме разработали обширную серию программ, направленных на улучшение здоровья рабочих. В нее вошли трепинговые программы, далеко превосходящие рамки проведения однократных лекций, включающие создание местных комитетов по охране здоровья, изменение обстановки на рабочих местах (например, удаление автоматов по продаже сигарет). Также все рабочие прошли опрос о риске для здоровья. По результатам опроса была составлена оценка риска, включившая в себя индивидуальные планы по улучшению здоровья. Эта оценка вошла в личные дела служащих.

После того как был окончен анализ потребности и принято решение о продолжении начатой работы, были разработаны детали программы и начата ее реализация. В ходе проведения программы было предпринят еще один вид оценки.

Наблюдение за программами — предварительная оценка

Программы нередко продолжаются довольно долгое время. С методологической точки зрения может казаться правильным ждать в течение года до проведения заключительной оценки эффективности программы, но что если уже в первый месяц ясно, что существуют проблемы, которые можно легко решить? Почему бы вместо ожидания окончания программы тщательно не понаблюдать за ней, пока

она находится в процессе реализации? Такое наблюдение называется **предварительной оценкой.** Согласно одному исследованию (Securest & Figueredo, 1993), она представляет собой наиболее часто встречаемый вид оценки программ.

В предварительную оценку входит несколько компонентов. Во-первых, она показывает, соответствует ли плану реализация программы. Предположим, к примеру, что местная организация, создавшая «горячую линию», решила разработать программу помощи детям, которые остаются одни после школы, так как их родители работают. Одна из очень важных задач — это сделать номер телефона горячей линии доступным и хорошо известным всем. В ходе предварительной оценки будет выявлено, помещена ли в надлежащее время реклама в газетах и были ли разосланы письма с наклейками, на которых указан телефонный номер горячей линии, как это было запланировано. Кроме того, можно выборочно обзвонить жителей и спросить, слышали ли они о новой горячей линии для детей. Нет никакого смысла проводить оценку эффективности программы, если люди даже не слышали о ней.

Вторая функция предварительной оценки — обеспечить получение данных о том, как используется данная программа. Исследователи позаимствовали название из области бухгалтерии и иногда называют такую процедуру **программным аудитом.** Так же как финансовый аудитор ищет случаи несоответствия того, как должен проводиться учет, с тем, как он в действительности проводится, аудитор программы изучает, соответствует ли реализуемая программа описанию, представленному в документации данной организации.

Заключительная часть предварительной оценки — это пилотажное исследование. Реализацию программы и начальные результаты можно оценить на небольшом примере, прежде чем проводить ее полномасштабное внедрение. Так и было сделано в ходе исследования в фирме «Дюпон». Пилотажное исследование на одном из заводов показало значимое снижение количество дней на больничном после введения программы по улучшению здоровья. Это воодушевило авторов программы и привело к ее реализации на других заводах (Bertera, 1990).

Оценка результатов — суммарная оценка

С политической точки зрения предварительная оценка менее опасна, чем **суммарная оценка,** представляющая собой общую оценку эффективности программы. Предварительная оценка направлена на усовершенствование программы и вряд ли поставит под вопрос само ее существование. Суммарная оценка вполне может это сделать. Если программа неэффективна, то зачем продолжать ее реализацию и платить ее руководителю и персоналу (понимаете, что я имел в виду, называя такую оценку «опасной»)? По мнению Сикреста и Фигередо (Sechrest & Figueredo, 1993):

> Суммарная оценка и даже причина ее проведения ставят под сомнение само существование организаций. Предварительная оценка, наоборот, просто отвечает на вопрос «Как можно сделать лучше?», практически не касаясь вопроса «Приносим ли мы вообще какую-нибудь пользу?» (р. 661).

Несмотря на эти политические трудности, суммарная оценка является ядром процесса оценки и важнейшей особенностью всех программ, имеющих федеральное финансирование. Любая организация, собирающаяся потратить средства, полученные от сбора налогов, на разработку программы, обязана показать, что деньги будут потрачены с пользой.

В действительности суммарная оценка подразумевает применение методов, с которыми вы уже успели познакомиться, а в особенности подходов, связанных с квазиэкспериментальными планами. Но иногда можно провести более точный эксперимент с использованием случайного распределения, особенно при оценке программы, на которую рассчитывают больше людей, чем существует возможностей. В таком случае случайное распределение, проводящееся в форме лотереи (случайные победители становятся участниками программы, а остальные записываются в лист ожидания), — это не просто методологический идеал, а единственная справедливая процедура.

Исследователи, проводящие оценку программы, иногда сталкиваются с вопросом о том, как объяснить то, что между экспериментальной и контрольной группами не обнаружено значимых различий. На языке статистики это означает «невозможность исключить H_0». Такой результат может оказаться весьма трудным для интерпретации. Вероятно, различий просто не существует, но всегда есть возможность совершения ошибки 2-го рода (см. главу 4), особенно если измерительные инструменты не слишком точны или надежны. Программа может иметь небольшой, но важный эффект, но в ходе анализа обнаружить его не удастся.

Хотя невозможность обнаружить различия может вызвать проблемы, большинство исследователей считают, что такой результат (особенно если его удалось воспроизвести) дает важную информацию, необходимую для принятия решения, особенно в случае прикладных исследований. Например, человек, выступающий за продолжение проведения новой программы, обязан показать, чем эта программа лучше других, уже существующих. Если не будут показаны различия между новой программой и уже существующей, то новую программу стоит прекратить, особенно если она требует большего финансирования. Как указали Итон и Сикрест (Yeaton & Sechrest, 1986), это может стать вопросом жизни и смерти. Они привели в пример исследование, сравнивающее виды лечения рака груди, различающиеся степенью хирургического вмешательства. Процент выживших женщин, перенесших минимальные операции, значимо не отличался от процента выживших из числа подвергшихся полной мастэктомии. Такой результат является примером «невозможности исключить H ». Стоит ли проигнорировать его патом основании, что не было обнаружено значимых различий? Конечно, нет. Если нужный результат достигается при менее травмирующей процедуре лечения, зачем делать более радикальные шаги? Конечно, уверенность в правильности решения, принятого на основании отсутствия различий, непосредственно зависит от регулярности возможного воспроизведения данного открытия. Рекомендации об использовании минимального хирургического вмешательства могут быть сделаны только после многократного подтверждения отсутствия различий между умеренными и радикальными процедурами.

«Невозможность исключить H_0» может помочь правильно оценить утверждения защитников новой программы. Человек, утверждающий, что обязательное тюремное заключение после третьего уголовного преступления заметно снизит число серьезных преступлений, обязан продемонстрировать наличие подобного эффекта. Результаты рассмотренного ранее исследования этого вопроса, проведенного по плану с прерванной временной последовательностью (Stolzenberg & D'Alessio, 1997), в ходе которого не было обнаружено никакого влияния программы «трех ударов» на уровень преступности, важны, несмотря на то что «различия не были обнаружены». Воспроизведение этих результатов привело бы к серьезным вопросам о значении закона об обязательном тюремном заключении.

Пример «невозможности исключить H_0» говорит о том, что отсутствие различий очень важно, когда при принятии решения учитываются возможные затраты на проведение программы. Таким образом, мы подошли к последнему виду оценки программ.

Оценка стоимости: анализ затрат и эффективности

Предположим, что исследователь занимается изучением вопроса здоровья и физической подготовки рабочих и проводит сравнение двух оздоровительных программ. Одна из них включает создание условий для физкультурных занятий в рабочее время, семинары, посвященные умению справляться со стрессом, и запрет курения. Вторая программа более обширная (и более дорогостоящая). Она включает оценку состояния каждого рабочего, разработку индивидуальных программ физических занятий и оплату таких процедур, как снижение кровяного давления и уровня холестерина. Программы проверяются на двух заводах; третий завод используется в качестве контрольной группы. Таким образом, используется план с неэквивалентной контрольной группой и не с одной, а двумя экспериментальными группами. Суммарная оценка не показала различий в улучшении здоровья рабочих между двумя экспериментальными группами, но по сравнению с контрольной обе экспериментальные группы продемонстрировали улучшение. Это значит, что обе программы оздоровления эффективны, но версия Шевроле так же хороша, как и Мерседес-версия. Если две программы приводят к одинаковым результатам, но различаются по стоимости, то к чему использовать более дорогую?

Данный пример с физическими занятиям на предприятии иллюстрирует один из видов **анализа затрат и эффективности:** изучение действительной стоимости программы и сравнение ее с эффективностью. Если две программы, имеющие одну цель, одинаково эффективны, но первая стоит наполовину меньше, чем вторая, то очевидно, что лучше использовать первую. Второй вид анализа стоимости проводится на этапе планирования программы. Оценка стоимости в самом начале проекта помогает определить пригодность данной программы и является основой для позднейшего сравнения планируемых и реальных затрат.

Исследование Эрфурта с соавторами (Erfurt, Foote & Heirich, 1992) иллюстрирует анализ затрат и эффективности. Исследователи сравнили несколько различа-

ющихся по стоимости «оздоровительных программ для рабочих» на четырех автомобильных заводах. Они обнаружили, что «в сравнении с общеобразовательными лекциями (рабочая площадка В) предоставление рабочим возможности для занятий спортом (площадка В) не привело к изменению степени риска для здоровья...» (р. 5). Так как стоимость программы на площадке *А* равнялась 17,68 долларов па рабочего, а на площадке *В* — 39,28 на рабочего, создание условий для занятий спортом было либо излишним (т. е. не оправдывало затрат), либо проведено неверно.

Сравнение затрат с результатами — процесс довольно сложный и нередко требует участия экономиста. Подробное обсуждение процедур для такого сравнения не входит в цели данной главы. Это мнение разделяет моя жена. Она по профессии экономист и находит мои зачаточные знания по этому вопросу весьма забавными. Основные понятия анализа затрат рассматриваются в главе 11 (Posavac & Carey, 1997).

Примечание по качественному анализу

В главе 3 были рассмотрены различия между количественным (численным) и качественным (числа не так важны) анализом. Хотя при оценке программ применяется в основном количественный анализ, качественный анализ также активно используется, особенно в первых трех описанных видах оценки. Так, при анализе необходимости количественные данные, собранные при опросе общественности, и данные переписи могут сочетаться с информацией, полученной в ходе подробных опросов ключевых информаторов и фокус-групп. При предварительной и суммарной оценке количественные данные могут быть дополнены качественным анализом опросов сотрудников и клиентов организации и непосредственным наблюдением работы программы. В целом, в исследованиях по оценке программ редко возникает вопрос о том, какой вид анализа лучше, качественный или количественный. Чаще всего исследователи используют (и высоко ценят) оба этих вида.

Как уже отмечалось в главе 5 при обсуждении внешней валидности, психологические исследования нередко критикуются за недостаточное внимание к решению практических проблем. Эта глава, посвященная прикладным исследованиям, должна показать, что подобная критика необоснованна. В действительности проблему применения и обобщения результатов исследований понимают все психологи, даже те, которые занимаются в основном фундаментальными исследованиями. История психологии показывает, что применение сделанных открытий — это основа американской психологии, хотя бы потому, что американцы не могут не думать о практической стороне дела. Поиск возможностей для практического применения так же неотъемлем от Америки, как яблочный пирог. В следующей главе вы познакомитесь с другим направлением психологических исследований — подробным изучением отдельных индивидуумов. Как вы уже знаете, корни прикладных исследований следует искать в работах первых психологов. Эксперименты с малым количеством участников (малым *hi)* также зародились на заре развития этой науки. Но прежде, чем переходить к изучению главы И, прочитайте вставку 10.3, где обобщаются некоторые этические проблемы, с которыми сталкиваются исследователи в ходе оценки программы.

ВСТАВКА 10.3

Этика — оценочные исследования и этика

Вне зависимости от того оценивается ли программа, реализуемая на предприятии или определенная правительственная программа, исследователи, занимающиеся ее оценкой, часто сталкиваются с этическими проблемами, которые никогда не встают перед психологами, работающими в лабораториях. Вот некоторые из них,

- *Осведомленное согласие.* Социальную помощь часто получают люди беспомощные. Когда их просят «добровольно поучаствовать» в исследовании и подписать документ об осведомленном согласии, они могут решить, что отказ «подписаться» будет означать потерю социальной поддержки. Исследователи должны совершать хорошо обдуманные действия, чтобы в такой ситуации убедить участников в ошибочности подобного мнения.
- *Сохранение конфиденциальности.* В некоторых исследованиях для сохранения конфиденциальности полученные от участников данные могут оставаться неподписанными. В других исследованиях необходимо знать, от кого именно получены данные. Например, исследователю может понадобиться вновь связаться с участниками, особенно если проводится лонгитюдное исследование. Или может потребоваться узнать, кто прошел опрос, чтобы связаться с теми, кто не прошел его. В таком случае следует разработать систему кодировки, чтобы скрыть имена участников. Иногда участники лонгитюдных исследований пользуются псевдонимами. Заполнив опросник, участник может анонимно послать его исследователю и сообщить о своем участии в отдельном письме (Sieber, 1998).
- *Воспринимаемая несправедливость.* Как было отмечено во вставке 7.2, некоторые люди могут возражать против участия в исследовании в качестве члена контрольной группы на том основании, что они не подвергаются предполагаемо благотворному воздействию. Хотя чаще всего члены контрольных групп подвергаются обычному воздействию, а не остаются вовсе без никакого, может возникнуть проблема контрольной группы. Например, если члены контрольной группы получат важную информацию о программе, в которой участвуют другие люди, может произойти «утечка» (см. вставку 8.2). Их возмущение по поводу того, что с другими «обращаются особенным образом», может серьезно повлиять на результаты исследования. Например, в исследовании, посвященном оценке изменений в работе угольной шахты, шахтеры из контрольной группы быстро начинают завидовать членам группы воздействия, которым, по их мнению, уделяют особое внимание и которые не должны так тяжело работать при той же зарплате (Blumberg & Pringle, 1983). Неприязнь распространилась и на самих исследователей. Рабочие из контрольной группы посчитали, что те находятся в сговоре с владельцем шахты с целью нарушить единство внутри профсоюза. Исследование в том виде, как оно было запланировано, пришлось прекратить.
- *Избегание конфликтов с заинтересованными лицами.* Заинтересованные лица - это люди, связанные с программой и материально заинтересованные в ее проведении, - заказчики, персонал и руководители программы. Исследователям, занимающимся оценкой программы, необходимо избегать возможных конфликтов с заинтересованными лицами. Это означает знать требования заинтересованных лиц и явным образом учитывать их на этапе планирования оценки. Например, исследователь может внести в контракт особые положения о том, в каком виде будет представлен отчет и каковы будут последствия получения различных видов результатов (Posavac & Carey, 1985). Говоря словами Кэмпбелла, четкие представления, составленные заранее, помогут превратить «директивную» администрацию в «экспериментирующую» (вставка 10.2).

Резюме

За пределами лаборатории

Задача прикладных исследований — пролить свет па причины возникновения и способы решения практических проблем. Но аналогично фундаментальным исследованиям их результаты важны для построения общих теорий о поведении (например, исследование, проведенное методом когнитивного интервью, расширяет наши знания о влиянии контекста на запоминание). Американские психологи всегда интересовались прикладными исследованиями, отчасти вследствие общественного давления, заставляющего их продемонстрировать, что психология — «новая», возникшая в конце XIX в. наука — способна быть полезной обществу. Прикладные исследования сталкиваются с этическими проблемами (например, обоснованное согласие) и проблемами внутренней валидности (например, неэквивалентность групп), но внешняя валидность у таких исследований всегда высока.

Квазиэкспериментальные планы

Исследования, в которых участники не могут быть случайно распределены по условиям, называются квазиэкспериментальными. Один из примеров — исследование с неэквивалентной контрольной группой. Обычно проводят сравнение изменений, произошедших от предварительного к заключительному тестированию в группе, подвергающейся некоторому воздействию, с подобными изменениями в контрольной группе, сформированной без использования случайного распределения. Эффект регрессии может затруднить интерпретацию результатов, если неэквивалентные группы уравниваются по оценкам предварительного тестирования. При использовании планов с временной последовательностью исследователи проводят несколько измерений до и после применения воздействия, которое необходимо оцепить. Исследования с прерванной временной последовательностью позволяют оценить влияние различных тенденций. Иногда к стандартному плану с временной последовательностью добавляют неэквивалентную контрольную группу, «переключающее» повторение или дополнительные зависимые переменные.

Оценка программ

Оценка программ — это один из разделов прикладной психологии. В ходе исследований по оценке программ собираются эмпирические данные об эффективности различных услуг, предоставляемых населению, или правительственных программ. Анализ необходимости выявляет, должна ли разрабатываться новая программа. Данные переписи, различных опросов и другие данные, полученные от населения, помогают оценить необходимость. Предварительная оценка показывает, проводится ли программа в соответствии с планом, а в ходе суммарной оценки оцениваются результаты исследования. Анализ затрат помогает определить, стоит ли польза, полученная от программы, затрат на ее реализацию. Исследование по оценке программ объединяет качественный и количественный подходы.

Задания для повторения

Выбор ответа

1. Прикладные исследования:
   1) как правило проводятся за пределами лабораторий;
   2) молодой подход в истории американской психологии;
   3) дают результаты, не подходящие для проверки общих теорий поведения;
   4) это синоним исследований по оценке программ.
2. Какие из перечисленных ниже проблем встречаются в прикладных исследованиях?
   1) невозможно получить осведомленное согласие участников;
   2) результаты редко оказываются полезными для построения, подтверждения или опровержения психологических теорий;
   3) низкая внутренняя валидность;
   4) результаты редко обладают внешней валидностью.
3. В плане с неэквивалентной контрольной группой:
   1) важно, чтобы обе группы получили одинаковые оценки предварительного тестирования;
   2) результаты позволяют исследователю оценить линейные тенденции;
   3) ключевым измерением зависимой переменной является оценка различий между результатами предварительного и заключительного тестирования;
   4) осложнение отбор x история можно исключить вследствие наличия контрольной группы.
4. Предположим, по причине большого количества травм у детей в штате Огайо решили ввести закон о ремнях безопасности для детей. Какое из следующих утверждений верно?
   1) скорее всего, количество травм возрастет в ближайшие год-два, так как тенденции не могут быстро измениться;
   2) для оценки эффективности данного закона лучше всего использовать план с неэквивалентной контрольной группой;
   3) сравнение со сходными штатами, в которых не действует подобный закон, поможет оценить возможный эффект регрессии;
   4) в такой ситуации качественный анализ гораздо эффективнее количественного.
5. Служба помощи семьям ввела новую программу по консультированию по вопросам задолженностей. Исследователь, занимающийся оценкой этой программы, периодически присутствует при таких обсуждениях, чтобы определить, соответствует ли происходящее плану. К какому виду относится такая процедура по оценке программы?
   1) анализ потребности;

2) суммарная оценка;
3) анализ затрат;
4) предварительная оценка.

Короткие эссе

1. Используя пример с когнитивными интервью, покажите: а) различия фундаментальных и прикладных исследований и б) что прикладные исследования связаны с разработкой теорий.
2. Опишите основные вклады Гуго Мюнстерберга в развитие прикладной психологии.
3. Опишите важнейшие особенности плана с неэквивалентной контрольной группой и объясните, почему рис. 10.3, в не позволяет заключить, что программа была успешной.
4. Первоначальная оценка программы «Рывок» показала, что успехи, которых добились участвующие в программе дети, были кратковременны: к третьему классу между детьми, принимающими участие в программе, и остальными не существовало никаких различий. Но такой результат мог быть вызван эффектом регрессии, который появился из-за того, что для формирования групп использовалась процедура уравнивания. Объясните данное явление.
5. Опишите пример, в котором оценивалась программа применения «тренинга эффективности для тренеров», и объясните, почему исследователи решили использовать не экспериментальный, а квазиэкспериментальный план.
6. Опишите важнейшие особенности плана с прерванной временной последовательностью и три разновидности основной процедуры, которые делают выводы исследования более основательными.
7. Опишите две количественные и две качественные процедуры, которые можно использовать при проведении анализа потребности.
8. Опишите различия между предварительной и суммарной оценками и объясните, почему организации предпочитают использовать первую.
9. Иногда в ходе исследований по оценке программы не удается обнаружить никаких различий. Объясните, почему такой результат не всегда плох.
10. Кратко опишите основные черты четырех видов исследований по оценке программ.
11. Кратко опишите этические проблемы, с которыми можно встретиться при проведении оценочных исследований.

Упражнения

Упражнение 10.1. Обнаружение факторов, угрожающих внутренней валидности

В квазиэкспериментальных исследованиях угроза внутренней валидности — явление весьма распространенное. Ниже приведен список факторов, угрожающих внутренней валидности, с которыми вы встретились в этой главе и главе 5. Для каждо-

го из описанных ниже гипотетических экспериментов определите, какой из этих факторов дает наиболее разумное альтернативное объяснение результатов.

Некоторые факторы, угрожающие внутренней валидности:

история	взросление
регрессия	отбор
истощение	отбор χ история

1. Директор колледжа озабочен низким процентом студентов, продолжающих обучение в колледже на втором курсе. Обычно их число было около 75%, но в начале учебного года в колледж вернулось лишь 60% бывших первокурсников. Директор ввел особую кураторскую программу, а когда на следующий год количество вернувшихся студентов составило 65%, похвалил ее эффективность.
2. Два соседних колледжа решили совместно провести оценку новой компьютеризированной системы обучения. Колледж *А* проводит программу, а колледж В использует уже существующие подходы к обучению. Когда исследование было проведено наполовину, колледж *В* заявил о своем банкротстве. Через год в колледже *А* повысился уровень компьютерной грамотности.
3. Двенадцать женщин, участвующих в программе домашних родов, сравнили со случайной выборкой других женщин, у которых роды проходили в больнице. Женщины из первой группы рожали среднем по 6 часов, а женщины из контрольной группы — в среднем по 9 часов.
4. Разработана шестинедельная программа по управлению экзаменационной тревожностью. В ней приняла участие выборка из студентов колледжа первого семестра. По окончании программы их уровень тревожности стал значительно ниже, чем он был вначале.
5. Преподаватель решил использовать новаторскую методику обучения, согласно которой в течение семестра студенты будут обучаться с наиболее подходящей для них скоростью. Курс содержит 10 разделов и студенты переходят к изучению раздела N, только изучив раздел N-1. По окончании изучения всех 10 разделов курс закончен и студент получает оценку А. Для 30 студентов, записавшихся в группу, получено следующее распределение оценок:

   A: 16 F:2 W: 12

   Преподаватель решил, что новая методика обучения не имела успеха.
6. Компания решила ввести свободный график работ. В январе была измерена производительность труда, затем в течение 6 месяцев использовалась новая программа, а в июне снова была измерена производительность. Был обнаружен рост производительности труда.

Упражнение 10.2. Интерпретирование исследований с неэквивалентными контрольными группами

Производитель роликовых подшипников владеет двумя заводами в штате Иллинойс и хочет проверить, можно ли снизить затраты на здравоохранение, если ввести программу охраны здоровья. Один завод (Е) выбран для проведения годичной экспериментальной программы, включающей оценку здоровья рабочих и разработку индивидуальных физкультурных программ. Второй завод (С) — это контрольная

группа. Уровень «отсутствий по болезни» операционально определен как количество дней в году, проведенных 100 рабочими на больничном. Он измеряется в начале и конце экспериментального года. Ниже представлены четыре набора результатов. Постройте график для каждого из наборов и определите, какой из них (возможно, ни один из них) наиболее явно говорит об эффективности программы. Для результатов, не подтверждающих эффективность программы, приведите альтернативные объяснения видимого улучшения показателей экспериментальной группы:

Результат 1.	E: предварительное тестирование = 125	Заключительное тестирование = 100
	C: заключительное тестирование = 125	Заключительное тестирование = 140
Результат 2.	E: предварительное тестирование = 100	Заключительное тестирование = 100
	C: заключительное тестирование = 125	Заключительное тестирование = 100
Результат 3.	E:предварительное тестирование = 125	Заключительное тестирование = 100
	C: заключительное тестирование = 110	Заключительное тестирование =110
Результат 4.	E: предварительное тестирование = 125	Заключительное тестирование = 100
	C: заключительное тестирование = 130	Заключительное тестирование = 105

Упражнение 10.3. Интерпретирование исследований с временными последовательностями

Представьте себе исследование с временной последовательностью, в ходе которого оценивается влияние закона об обязательном ношении шлема на количество травм головы у хоккеистов. На следующий год после введения закона количество травм значимо снизилось по сравнению с предыдущим годом. Постройте три графика с временными последовательностями, один для каждого из следующих результатов:

1. Закон о ношении шлемов был эффективен.
2. Закон о ношении шлемов был эффективен, но его влияние было кратковременным.
3. Закон о ношении шлемов был неэффективен. Видимое снижение количества травм, скорее всего, связано с эффектом регрессии.
4. Закон о ношении шлемов был неэффективен. Видимое снижение количества травм, по-видимому, отражает общую тенденцию к снижению уровня насилия в спорте.

Упражнение 10.4. Планирование оценки необходимости

Вы — глава психологического отделения из пяти человек в гуманитарном университете искусств. Однажды декан говорит вам: «Почему бы вам не разработать программу для получения степени магистра по консультативной психологии?» В колледже уже действуют магистерские программы по администрированию в бизнесе и по лечебной физкультуре. Поскольку вы прочитали эту главу, вы отвечаете, что необходимо провести тщательный анализ потребности. Декан соглашается и даже выделяет небольшую сумму для реализации проекта. Опишите факторы, которые необходимо рассмотреть, прежде чем вводить программу получения новой степени, и методы, которые вы будете использовать при оценке потребности.

ГЛАВА 11

Планы с малым *N*

Обзор задач главы

Планы, которые мы до сих пор рассматривали, были связаны с исследованием относительно больших групп участников, изучением особых методологических проблем, таких как создание эквивалентных групп или устранение эффекта последовательности, вычислением значения среднего арифметического, проведением заключительного статистического анализа (например, методом *ANOVA)* и построением общих выводов о влиянии независимых переменных. В данной главе вы встретитесь с совершенно новым типом исследований. Планы таких исследований часто называют «планами с одним субъектом», так как поведение каждого из субъектов исследования рассматривается индивидуально. Также их называют «планами с малым *N*», поскольку такие исследования иногда включают несколько участников. Полученные от испытуемых данные *можно* статистически обобщить, но чаще всего информация, получаемая от новых участников, описывается индивидуально и рассматривается как повторение исследования. Изучив эту главу, вы:

- сможете описать несколько примеров классических исследований отдельных индивидуумов или исследований с малым количеством участников, в которых изучение новых испытуемых проводится с целью повторения исследования;
- сможете объяснить, почему обобщение данных, полученных от большого числа участников, может привести к неверным выводам о поведении, и описать практические основания для проведения исследований с малым *N;*
- сумеете описать философию проведения исследований Б. Ф. Скиннера — экспериментальный анализ поведения;
- научитесь описывать основные компоненты плана с одним субъектом;
- сможете объяснить логику плана с отменой А-В-А-В;
- сумеете объяснить логику плана с несколькими базовыми уровнями и описать три вида процедуры с несколькими базовыми уровнями;
- сумеете объяснить логику плана с изменяющимся критерием в связи с оперантным понятием формирования;

- сможете описать другие планы с одним субъектом (например, план с чередующимся воздействием);
- сможете описать критику исследований с малым *N* со стороны оперантного направления;
- сможете описать цели психофизических исследований и сравнить три наиболее распространенных метода психофизики.

Стратегию проведения исследований с малым *N* чаще всего связывают с именем Б. Ф. Скиннера, с которым вы уже встречались в главе 1. Но важно знать, что Скиннер не был первым ученым, сконцентрировавшим свое внимание на исследовании отдельных субъектов. Планы с малым Nимеют довольно длинную историю — уже первые психологи-экспериментаторы использовали этот подход в своей работе.

Первые исследования по психологии и малое *N*

Когда во второй половине XIX в. психология сформировалась как новая наука, основы статистического анализа еще только закладывались. Гальтон только начинал разрабатывать теорию корреляций, а такие методы, как дисперсионный анализ *(ANOVA)* еще не существовали. Широкое применение планы с большим N *и* методы статистической обработки информации получили только после того как в 30-е гг. вышла в свет работа Фишера, посвященная дисперсионному анализу (см. главу 8 вставку 8.3). До тех пор господствовали планы с малым *N*.

Некоторые из первых психологов использовали наименьшее из возможных *N*. они изучали свое собственное поведение или поведение отдельных индивидуумов. В главе 7 (вставка 7.3) вы познакомились с одним из самых известных примеров — исчерпывающим исследованием Германом Эббингхаузом своей способности запоминать и воспроизводить списки слогов. Другой пример — исследование развития ребенка, проведенное Чарльзом Дарвином. Ученый вел подробный дневник детства своего собственного сына. Дневник был опубликован под названием «Биографические наброски об одном ребенке» в британском журнале «Майнд» в 1877 г. (Goodwin, 1999). Третий пример — печально известный эксперимент Уотсона и Рейнер с маленьким Альбертом (Watson & Rayner, 1920) (см. главу 2, вставку 2.1).

В лаборатории Вундта в Лейпциге ведущей стратегией исследований также были планы с малым *N*. Аспирантам, стремившимся получить докторскую степень, выдавались задания для исследований, которые обычно занимали около года. Чаще всего эти исследования включали очень малое количество участников, а сам исследователь нередко входил в их число. Другими испытуемыми обычно были другие аспиранты. Например, ряд диссертационных исследований Джеймса Кеттелла, посвященных скорости реакции, включал данные лишь двух испытуемых — самого Кеттелла и его друга и коллеги-аспиранта Густава Бергера (Sokal, 1981). Очевидно, что существующее сегодня разделение ролей (и статуса) между Экспериментатором с большой буквы и субъектом с маленькой в то время не существовало. По сути, хотя в 90-х гг. XIX в. участников иногда называли «субъектами», более правильно было бы использовать для них термин «наблюдатели», что предполага-

ет более высокий статус и более активную роль по сравнению с субъектами. Вопрос о том, какой термин использовать, субъект или наблюдатель, был поднят только в 1930 г. (Danziger, 1985).

Первые психологи-экспериментаторы иногда проводили грубое обобщение данных (например, находили среднее арифметическое), полученных от нескольких наблюдателей, но чаще всего они составляли отчеты для каждого участника отдельно. Эту стратегию отлично иллюстрирует исследование, проведенное в лаборатории Университета Кларк в 90-х гг. XIX в. Оно было посвящено так называемому «лицевому зрению» — способности различать близлежащие объекты даже при отсутствии возможности их увидеть. Одно время считалось, что у слепых эта способность развивается как особое чувство, компенсирующее отсутствующее зрение. Но Флетчер Б. Дресслер (Dressier, 1893) смог показать, что подобное чувство связано скорее со слухом, чем со зрением.

На рис. 11.1 приведена фотография экспериментальной ситуации. Она входит в серию фотографий, отображающих ход исследовательского процесса в Университете Кларк в 1892 г. Как вы видите, человек с завязанными глазами сидит рядом с панелью, составленной из четырех квадратов со стороной около 30 см. Слева направо квадраты либо пустые, либо закрыты: а) деревянной решеткой, б) сплошной деревянной панелью или в) проволочной сеткой. Панель подвешена к потолку и может передвигаться экспериментатором (на фотографии изображен Дресслер) так, чтобы каждый из квадратов мог быть помещен перед лицом испытуемого. Задание состоит в том, чтобы определить, какая именно из поверхностей находится перед лицом. Испытуемыми были Дресслер и два других аспиранта.

Примечательно, что, как показано в табл. 11.1, сопровождавшей статью об исследовании, все три участника научились различать пары поверхностей. В таблице представлено количество правильных (П) или неправильных (Н) ответов. Например, при сравнении решетчатой поверхности со сплошной Ф. Б. Д. (F. B. D.) (угадайте, кто это) ответил правильно 69 раз и ошибся только один раз, когда правильный ответ был «решетка», и ответил правильно 70 из 74 раз при правильном ответе «сплошная». Два других участника показали аналогичные результаты.

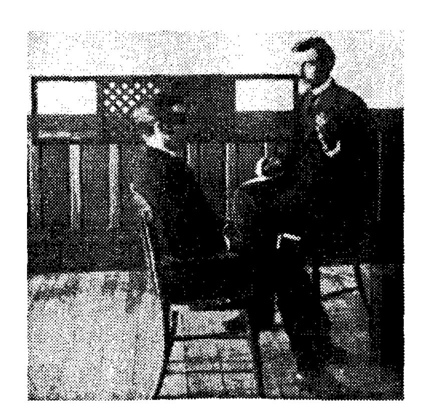


Рис. 11.1. Устройство Дресслера для изучения лицевого зрения

Таблица 11.1

Данные исследования лицевого зрения Дресслера

Испытуемый	Пустая и решетка				Решетка и сплошная				Сплошная и сетка			
	П	Н	П	Н	П	Н	П	Н	II	Н	**П**	**Н**
Дж. А. Б.	65	15	59	25	58	2	56	0	45	0	46	2
О.К.	72	47	74	46	33	13	28	14	21	4	14	9
Ф.Б.Д.	53	24	58	17	69	1	70	4	73	0	77	2

Обратите внимание, что хотя данные представлены для всех трех участников, статистическое обобщение информации не сделано. Это связано с задачей исследования, которая заключалась в том, чтобы показать, что явление обязательно возникает в случае каждого испытуемого, а не в случае обобщенного среднего испытуемого. Дресслер изучил двух дополнительных испытуемых, чтобы дважды *повторить* первоначальное открытие. Сегодня такой подход часто используют в исследованиях с малым *N*.

Говорят ли результаты исследования Дресслера о том, что существует лицевое зрение как отдельное чувство? Нет. Как хороший психолог-исследователь, Дресслер попытался найти более простое объяснение этого явления и исключить (фальсифицировать) существование особого лицевого чувства. Ему удалось достичь желаемого, когда он внес в процедуру исследования небольшое изменение — заткнул уши испытуемым. Результат был налицо: их «способность различать [панели] была полностью потеряна» (Dressier, 1893, p. 349). Так лицевое зрение обернулось способностью устанавливать легкие различия в отражаемых звуковых волнах[1].

Исследования, подобные проведенному Дресслером, в которых рассматриваются данные одного или нескольких (но небольшого числа) участников, были широко распространены на заре экспериментальной психологии, но это не значит, что исследования с большим N совсем не проводились. Например, их можно найти в области психологии образования и в исследованиях, посвященных развитию детей (Danziger, 1985). Подобные исследования ставили перед собой такие эмпирические вопросы, как «Чего боятся школьники младших классов?» и обобщали результаты анкет, заполненных сотнями детей (например, Hall, 1893). Но, как было отмечено выше, только с 30-х гг, после того как вышла работа сэра Рональда Фишера, психологи стали собирать данные у большого количества участников и проводить описательный и заключительный статистический анализ, как это делается сегодня.

Один пример, описанный во вставке 11.1 и посвященный использованию плана с малым *N,* стоит обсудить более подробно. Поскольку это исследование, выполненное 100 лет назад, является важным этапом в работе Б. Ф. Скиннера по оперантному обусловливанию и предвещает появление бихевиоризма, оно по праву считается классическим. Также оно показывает, что хорошее научное исследование может быть проведено с весьма скромным бюджетом и минимальными способностями экспериментатора к конструированию приборов.

[1] Получив такие результаты, что бы вы сделали дальше, будучи экспериментатором? Что если, к примеру, изменить расстояние между панелью и испытуемым?

ВСТАВКА 11.1 " ~

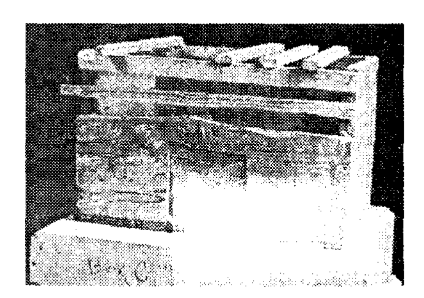
Классические исследования — кошки в ящиках-головоломках

ЭдЕард Л. Торндайк (1874-1959) сделал отдельную карьеру в области педагогической психологии. Но лучше ьсего помнят его исследование, проведенное в рамках докторской диссертации, посвященное изучению способности кошек выбираться из ящиков-головоломок (Thorndike, 1898). Это исследование важно по нескольким причинам: оно показывает, что первые психологи занимались подробным изучением отдельных испытуемых, является хорошим примером использования простых объяснений поведения и образцом того вида исследований, которые проложили дорогу бихевиоризму, особенно скиннеровской его разновидности.

Исследование отдельных кошек

Для изучения способности кошек к обучению Торндайк построил пятнадцать ящиков-головоломок, каждый из которых имел свою собственную стратегию прохождения. Историк Джон Бернем (Burnham, 1972) совершенно случайно обнаружил фотографии этих ящиков в бумагах Роберта Йеркса. Две из них представлены на рис. 11.2. Очевидно, что Торндайк обладал не слишком высокими способностями к конструированию приборов. В действительности его технические способности были настолько низкими, что он не смог научиться водить машину (Hothersall, 1990). Это дает нам важный урок: значительную исследовательскую работу можно проводить, не имея сложного оборудования. Идея исследования гораздо важнее внешнего лоска.

Кошки изучались по отдельности, и полученные результаты были изложены также отдельно для каждой кошки. Кошки учились выбираться из ящиков в процессе эксперимента, названного Торндайком (1911/2000) «пробы и ошибки со случайным успехом» (р. 150) и в соответствии с явлением, названным автором «законом эффекта» (р. 244). Сначала действия котов были случайными и лишь иногда, также случайно, успешными. Успешное поведение обычно повторялось («отпечатывалось», как выразился Торндайк), а неуспешное постепенно прекращалось («подавлялось»). Таким образом, «эффект» успешного поведения состоял в увеличении вероятности его появления при следующей попытке. На рис. 11.3 показан прогресс в действиях одной из кошек Торндайка (N10, из ящика C)[1].



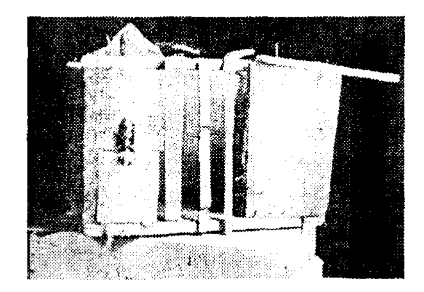


Рис. 11.2. Два из пятнадцати ящиков-головоломок (C и D), построенных и использованных Торндайком !? классическом исследовании поведения кошек

[1] Торндайк не разметил ось *X* на этом графике, а лишь укапал, когда между упорядоченными попытками прошло значительное количество времени. На рис. 10.3, к примеру, никак не обозначенная вертикальная линия означает день, «2» означает 2 дня, «1ч» — это час, а «78» — 78 часов.

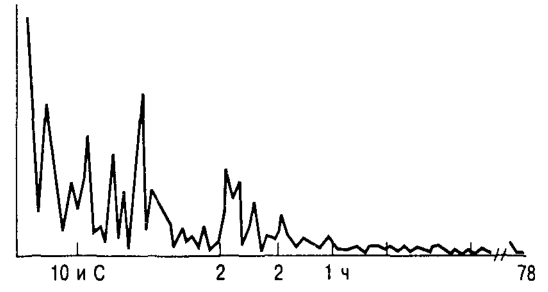


Рис. 11.3. Запись изучения кошкой №10 пути к выходу из ящика С

Использование простых объяснений

Открытый Торндайком закон эффекта поставил под вопрос распространенное мнение об умственных способностях животных и дал более простое объяснение их способности к решению задач. Ученый указал, что не следует приписывать животным способность к мышлению, если их поведение можно объяснить с точки зрения более простых процессов (т. е. применением базового метода проб и ошибок). Торндайк не признавал подхода, некритично приписывающего животным наличие мыслительных процессов. Такое отношение выразил Ллойд Морган (см. главу 3) в своем известном принципе простоты, уже сформулированном к тому времени. Торндайк был знаком с работой Моргана и, возможно, слышал о призыве англичанина к изучению животных, который он сделал во время своего посещения Гарварда в 1896 г. (Joncich, 1968).

В ожидании Скиннера

И последнее, что стоит отметить по поводу исследования Торндайка с ящиками-головоломками, это то, что оно представляет экспериментальный подход к исследованию процесса научения, проложивший дорогу другим исследователям поведения. Оно также дает образец процесса научения, впоследствии принявшего вид скиннеровского анализа поведения, с которым вы вскоре познакомитесь. Скиннер (Skinner, 1953) высказал свою признательность Торндайку, сказав, что работа последнего находится в числе «первых попыток изучения изменений, вызываемых последовательностью поведенческих актов» (р. 59).

Причины использования планов с малым *N*

Несмотря на популярность в современной психологии планов с большим *N,* исследования с одним или несколькими испытуемыми внесли и продолжают вносить важный вклад в развитие знаний о поведении. Как вы вскоре узнаете, такие изыскания охватывают полный диапазон от лабораторных до полевых и от фундаментальных до прикладных исследований. Есть несколько причин проведения исследований с малым *N.*

Дезориентирующие результаты обобщения данных

Обобщение данных, полученных на больших группах людей, иногда дает результаты, которые нельзя применить ни к одному из участников. Это значит, что результаты не обладают так называемой **индивидуальной валидностью** *(individual-subject validity)* (Dermer & Hoch, 1999). Индивидуальная валидность указывает, до

какой степени общие результаты исследования можно применить к отдельным участникам. Отсутствие такой валидности в исследованиях с большим количеством участников явилось лейтмотивом работы Сидмэна «Тактика научных исследований» (Sidman, *Tactics of Scientific Research,* 1960), которую сторонники исследований с малым Досчитают классическим руководством по методологии данного вида исследований. Поскольку средние показатели группы не отражают индивидуальных различий между испытуемыми, Сидмэн указал, что «групповые данные часто описывают процессы или функциональные зависимости, не являющиеся валидными для конкретных людей» (р. 274).

Но более важным, по мнению Сидмэна, является то, что усредненные данные могут свидетельствовать в поддержку теории *X,* тогда как они не должны этого делать. Рассмотрим один пример из материалов эксперимента по научению детей понятиям. Детям показывали большие наборы парных стимулов и просили определить, какой из стимулов в каждой паре использован верно. Если ребенок делал правильный выбор, то он получал вознаграждение. В качестве стимулов использовались простые геометрические фигуры. На рис. 11.4 показано, какие пары стимулов можно использовать для проведения семи попыток. Знак плюс говорит, какая из фигур в каждой паре получает подкрепление (к примеру, ребенку выдают конфету). Как показано на рисунке, стимулы различаются по форме (треугольники, квадраты или круги), цвету (красный или зеленый) и положению (слева или справа). Правильное понятие в данном примере — это «красный». Набрав большое количество конфет, ребенок должен понять, что форма и положение не играют в выборе никакой роли. Задание считается выполненным, когда ребенок достигает определенной «критериальной» оценки, например дает 10 последовательных правильных ответов.

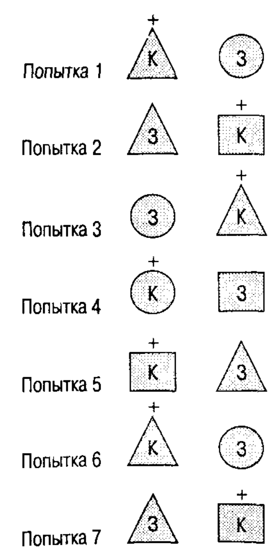


Рис. 11.4. Стандартные стимулы, используемые при исследовании способности детей к научению дифференциации

В литературе, посвященной научению понятиям, очень давно продолжается дискуссия о том, какой способ научения используется при выполнении данного задания (Manis, 1971, р. 64-68). Согласно теории «непрерывности», научение — это процесс постепенного накопления «силы привычки». Каждая подкрепленная попытка усиливает стремление реагировать на значимый показатель и ослабляет реакцию на другие показатели. График этого гипотетического нарастающего процесса научения показан на рис. 11.5, а. Теория «прерывности», в свою очередь, говорит о том, что при первых попытках испытуемые активно проверяют различные «гипотезы». В процессе поиска правильной гипотезы их действия носят случайный характер (уровень 50%), но когда она найдена, действия приобретают 100% -ную точность и сохраняют ее. С точки зрения теории прерывности график должен выглядеть, как показано на рис. 11.5, б.

История этого вопроса длинна и сложна, а общий вывод зависит от многих условий, в частности от способа обработки данных. Если данные, полученные от большого количества участников, сгруппировать и построить общий график, то результаты действительно будут похожи на рис. 11.5, а, отражающий теорию непрерывности. Однако если более пристально изучить действия отдельных участников, особенно при выполнении трудных заданий, то можно построить график, напоминающий рис. 11.5, б, что будет поддержкой теории прерывности (Osier & Trautman, 1961). Изучение действий участников незадолго до того, как найдено решение показывает, что точность достигает 50% (см., напр., Trabasso, 1963). После достижения критериальной оценки действия становятся безошибочными. Таким образом, участники действуют на уровне случайности до тех пор, пока не находят правильного решения, после чего качество выполнения задания значительно улучшается. Так как же индивидуальные действия, отображенные на рис. 11.5, б, превращаются в их изображение на рис. 11.5, а после обобщения данных?

Ключевым показателем здесь является время, требующееся каждому ребенку для нахождения правильного ответа: одни справляются заданием быстро, а другим требуется больше времени. Эта ситуация отображена на рис. 11.6. Как вы видите, ряд индивидуальных кривых при объединении дает гладкую кривую, показанную на рис. 11.5, а. Этот пример ясно показывает, как обобщенные данные могут создать картину результата, которая не подтверждается поведением отдельных участников. Поэтому исследователи, использующие планы с большим *N*, особенно при изучении вопросов научения, должны изучать индивидуальные данные, чтобы понять, соответствуют ли они обобщенным результатам.

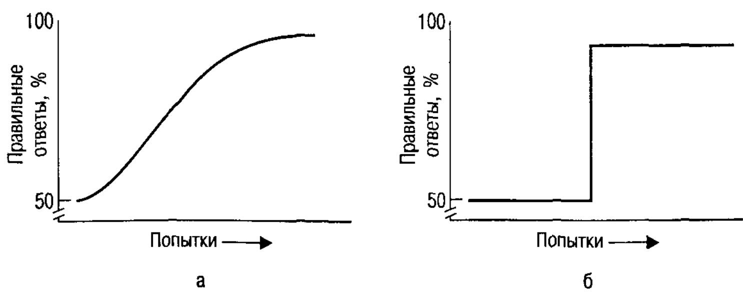


Рис. 11.5. Данные концептуального научения с точки зрения (а) теории непрерывности и (б) теории прерывности

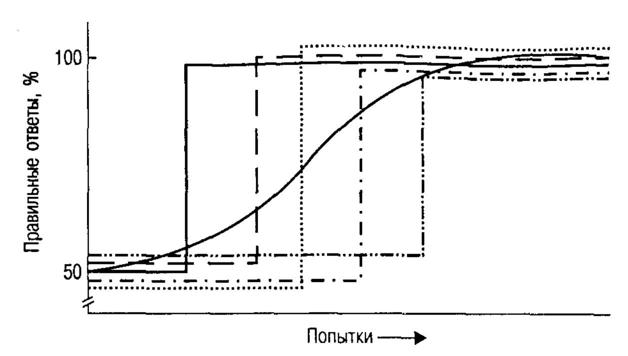


Рис. 11.6. Обобщение данных для отдельных детей в эксперименте по концептуальному научению дает гладкую, но неверную кривую

Практические проблемы планов с большим *N*

Когда потенциальных участников очень мало или их трудно найти, необходимо использовать план с малым *N*. Особенно часто такая ситуация возникает в клинической психологии, когда, к примеру, исследователь хочет изучить людей с редким заболеванием, и в психологии познания при изучении людей с необычными способностями к запоминанию (пример такого человека описан во вставке 12.3). Сходные проблемы появляются в некоторых исследованиях животных, особенно при использовании хирургического вмешательства, ведь хирургическое оборудование недешево, а применяемые процедуры отнимают много времени. Само содержание группы животных обходится довольно дорого, если учесть стоимость урегулирования вопросов с защитниками прав животных, включая затраты на систему безопасности. Животных некоторых видов бывает трудно получить, они могут чрезмерно дорого стоить или требовать длительного обучения. Например, для обучения шимпанзе и других человекообразных обезьян языку жестов требуется потратить сотни часов на каждое животное, поэтому исследования, включающие эту процедуру, обычно продолжаются многие годы. В ходе одного исследования языку жестов обучали гориллу (Patterson & Linden, 1981). К 10 годам обезьяна знала более 400 знаков, а началось исследование, когда ей только исполнился год.

Экспериментальный анализ поведения

Таким образом, результаты использования планов с большим N иногда не способны представить поведение отдельных испытуемых, и поэтому их нельзя использовать даже при большом желании. Но есть и философские причины предпочтения планов с малым *N*. Лучше всего их сформулировал Б. Ф. Скиннер (1904-1990) — самый известный защитник этого направления исследований. Как вы должны помнить из главы 4, при опросе историков и заведующих кафедр психологии Скиннер был признан самым выдающимся психологом современности (Korn, Davis & Davis, 1991).

Скиннер был убежден в том, что для достижения психологией своих целей, заключающихся в предсказании поведения и управлении им, необходимо интенсив-

но исследовать отдельные организмы и выводить общие правила и законы только после всестороннего изучения отдельных случаев. Таким образом, психология должна быть индуктивной наукой, переходящей от частных законов поведения к общим. Однажды Скиннер сказал, что исследователь должен «изучать одну крысу на протяжении тысячи часов», а не «тысячу крыс по часу каждую или сто крыс по десять часов каждую» (Skinner, 1966, р. 21). Задача состоит в снижении случайной изменчивости путем достижения строгого контроля за экспериментальными условиями, воздействующими на отдельных испытуемых. Как писал Скиннер: «Я получил путеводную нить от Павлова: контролируй условия и получишь точный результат» (Skinner, 1956, р. 223). Он назвал свою систему «экспериментальным анализом поведения». Чтобы получить более подробную информацию об идеях Скиннера, вам потребуется обратиться к другим источникам (например, Skinner, 1953), но самые важные из них, дающие философское обоснование исследованиям в области экспериментального анализа поведения, вы найдете в этой книге.

Оперантное обусловливание

Скиннер более всего известен своими исследованиями, посвященными **оперантному обусловливанию** — «процессу, при котором частота появления определенных поведенческих актов изменяется последствиями данного поведения» (Reynolds, 1968, р. 1). Это значит, что когда поведение появляется в конкретной ситуации, оно вызывает определенные последствия. Если это последствия положительные, то поведение будет иметь тенденцию к возобновлению, когда индивидуум окажется в такой же ситуации. Негативные последствия, в свою очередь, снижают вероятность появления определенного поведения в будущем. Если вспышка раздражения у ребенка «сработала» (т. е. привела к получению игрушки), возникнет тенденция к ее повторению, а если «не сработала», то такой тенденции не возникнет. Обратите внимание, что в определении используется выражение «частота появления». Скиннер считал, что в экспериментальном анализе поведения единственной зависимой переменной, стоящей изучения, является **частота реакции.** Если целью психологии является предсказание поведения и управление им, а по Скиннеру, это единственно важная цель, тогда нас будет интересовать только, появляется или нет определенный тип поведения и как часто он появляется в единицу времени.

Таким образом, по Скиннеру, характерное для нас поведение управляется окружающими условиями нашей жизни. По мнению этого ученого, чтобы предсказать поведение или управлять им, все что требуется — это «определить три вещи: 1) ситуацию, в которой возникает реакция, 2) саму реакцию и 3) подкрепляющие последствия. Взаимосвязь между ними — это ситуативное подкрепление» (Skinner, 1969, р. 7).

В лабораториях оперантное обусловливание чаще всего изучают с помощью прибора, называемого «проблемным ящиком Скиннера». На рис. 11.7 показана типичная камера для исследования крыс. Крыса нажимает выходящий из стены рычаг. Положительным последствием является попадание в кормушку пищи, что увеличивает вероятность повторного нажатия рычага крысой. Отрицательным последствием может быть короткий электрический разряд, подаваемый на решетку пола и уменьшающий вероятность нажатия рычага в будущем.

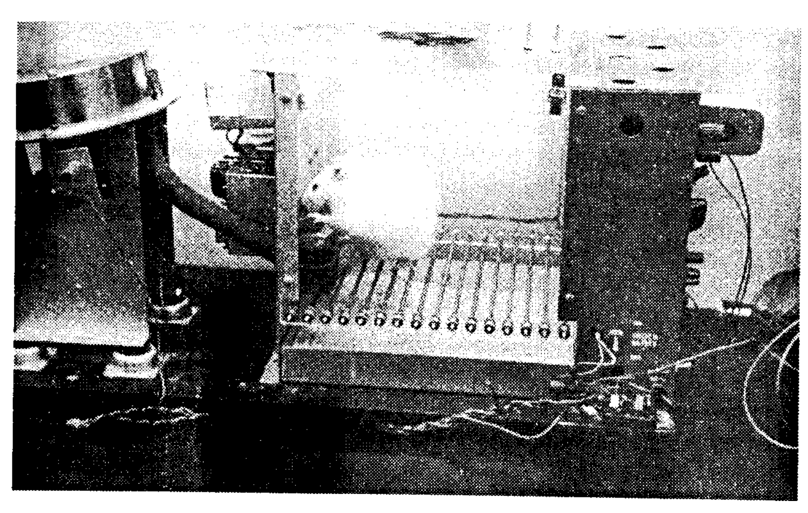


Рис. 11.7. Оперантная камера с кумулятивной записью для изучения крыс

Когда рефлекс нажимания рычага выработан, можно начать управлять им с помощью стимулов окружения, таких как лампочка, расположенная прямо над рычагом. Если кусочки еды появляются после нажатия рычага только при горящей лампочке, животное быстро научается проводить простую дифференциацию: нажимать при включенной лампочке и не нажимать при выключенной. На скиннеровском языке условий подкрепления зажженный в камере свет означает «ситуацию, в которой возникла реакция», «сама реакция» — это нажатие рычага, а кусочки еды — это «подкрепляющие последствия».

Частота нажатий рычага регистрируется прибором, называемым **«кумулятивным записывающим устройством».** На рис. 11.8, а изображен такой прибор в действии. Бумага поступает в прибор с постоянной скоростью, что позволяет отмечать время на оси *X*. Перо самописца передвигается но бумаге на определенное расстояние каждый раз, когда животное нажимает на рычаг. Если за реакцией следует подкрепление, перо чертит короткую вертикальную линию. Когда перо доходит до края бумаги, оно возвращается на исходную позицию и начинает процесс сначала. Второе перо отмечает такие события, как включение и выключение света. Частоту реакции легко оценить по наклону появившейся кривой. Кривая на рис. 11.8, б говорит о том, что крыса сначала нажимала рычаг очень часто (вероятно, сигнальный свет в ящике был включен и нажатие рычага приводило к получению еды), но во второй половине записи она вряд ли вообще прикасалась к рычагу (вероятно, лампочка не горела).

Классическую иллюстрацию оперантного подхода и его индуктивной философии исследований можно найти в энциклопедическом труде «Режимы подкрепления» (Fester & Skinner, Schedules of Reinforcement, 1957). Значение этой работы для экспериментального анализа поведения раскрыто во вставке 11.2. Прочитайте ее, прежде чем переходить к изучению графиков подкрепления, обсуждаемых в примере 27.

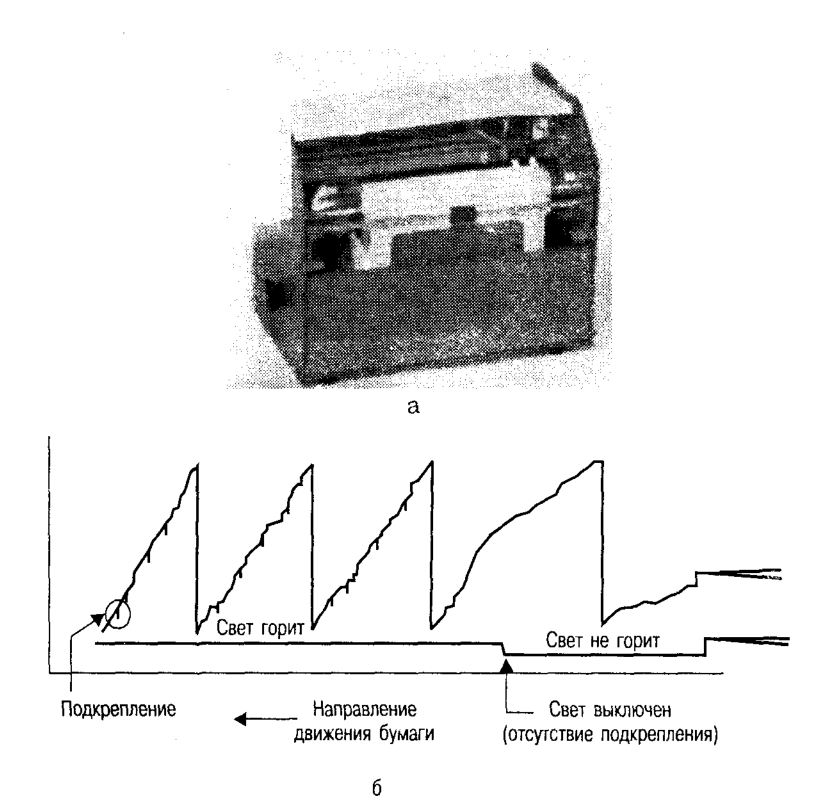


Рис. 11.8. а) Общий вид кумулятивного записывающего устройства и б) гипотетическая кумулятивная запись, показывающая и высокую, и низкую скорости реакции

Основной источник информации о фундаментальных исследованиях оперантного обусловливания — это «Журнал экспериментального анализа поведения». Ниже приведены заголовки, которые покажут вам, какого рода статьи встречаются в этом журнале:

«Влияние режимов с постоянным и переменным номером подкрепляемой реакции на изменчивость поведения людей» (Tatham, Wanchiscn, & Heneline, 1993);

«Влияние задержки подкрепления па степень вокализации у детей» (Reeve, Reeve, Brown, Brown, & Poulson, 1992);

«История обусловливания с низкой скоростью реакции и реакции с фиксированным временным интервалом у крыс» (LcFrancois & Metzger).

Пример 27. Экспериментальный анализ

Третье из перечисленных выше исследований представляет собой превосходный пример фундаментального исследования в области экспериментального анализа поведения. Лефрансуа и Мецгер (LeFrancois & Metzger, 1993) заинтересовались одной из важнейших проблем бихевиоризма: как предшествующее подкрепление влияет на текущее поведение? Все бихевнористы согласны с тем, что для прогно-

зирования поведения необходимо знать историю научения испытуемого, но их мнения расходятся, когда требуется ответить на вопрос о том, что *в первую очередь* влияет на актуальное поведение, более поздние условия подкрепления или более ранние. Если поведение адаптивно, то следует ожидать, что более поздние события будут иметь наибольший эффект. Однако можно возразить, что влияние особо мощного графика подкрепления, использованного в раннем возрасте, может рас-. пространяться и па текущее время.

ВСТАВКА 11.2

Классические исследования – режимы подкрепления

«Режимы подкрепления» (Ferster & Skinner, 1957) - это не та книга, которую стоит читать. Лучше держать ее на полке, и когда вас спросят, как режим с «постоянным номером подкрепляемой реакции» влияет на поведение, вы откроете главу о постоянном номере реакции и скажете: «Вот, посмотри сам». 739-страничный труд - это скорее «атлас» (термин Скиннера), чем книга для летнего чтения.

Написанная в соавторстве со студентом и коллегой Скиннера Чарльзом Фестером книга представляет собой выдающийся труд, возможно, даже важнейшее изыскание Скиннера в области психологических исследований и прекрасный пример *индуктивного* научного подхода. Работа наполнена многочисленными примерами описаний и кумулятивными записями, иллюстрирующими десятки различных видов условий подкрепления. Следующее описание процесса создания книги дает представление об этике работы Скиннера: «Тысячи футов кумулятивных записей представляли тысячи часов данных... Мы работали систематически. Мы брали протокол и пачку кумулятивных записей, диктовали отчет об эксперименте, отбирали пояснительные записи и составляли из них рисунки. В конце мы получили более 1000 рисунков, 921 из которых вошли в книгу» (Skinner, 1984, р. 109).

Кумулятивные записи (921!) в книге фестера и Скиннера показывают, как различные режимы подкрепления влияют на поведение. Считалось, что режим действует, «если окружающая среда давала подкрепление отдельным, но не всем, реакциям организма» (Reynolds, 1968, р. 59). Некоторые поведенческие акты подкреплялись каждый раз, когда они возникали, а другие (например, азарт) - лишь изредка. Кроме того, частичное подкрепление рождало поведение, очень устойчивое в отношении затухания (тот же азарт).

Из курса по общей психологии (глава, посвященная научению) вы, возможно, помните о двух распространенных типах режимов подкрепления: «с временным интервалом» и режиме «с номером реакции». При режиме с постоянным временным интервалом поведение подкрепляется только после того, как прошло определенное количество времени. При режиме с переменным временным интервалом поведение подкрепляется также по прошествии определенного времени, но время варьируется от подкрепления к подкреплению. Режимы с номером реакции учитывают не время, а количество поведенческих актов. При режиме с постоянным номером подкрепляемой реакции у индивидуума подкрепляется каждая *п-я* реакция. При режиме с переменным номером подкрепляемой реакции подкрепление наступает после определенного количества реакций, которое изменяется от подкрепления к подкреплению. Режимы с номером реакции дают относительно высокую скорость реакции, так как количество получаемых подкреплений зависит от частоты появления поведенческих актов. Режимы с временным интервалом обычно дают низкую скорость реакции, но среди них есть и исключения, о которых вы узнаете, изучив пример 27.

Лефрансуа и Мецгер в ходе оперантного эксперимента с шестью крысами исследовали обе эти возможности. Сначала всех крыс тренировали по режиму ДПН-20 с. «ДПН» означает «дифференциальное подкрепление низкой скорости реакции», при котором животные получают подкрепление нажатия рычага, *только* если перед этим прошло 20 секунд, за которые рычаг не нажимался. Затем половина крыс перешли на режим с постоянным временным интервалом (ПВИ), а половина — на режим с постоянным номером подкрепляемой реакции (ПНПР). В завершение для второй группы также использовался режим ПВИ. Таким образом, были получены следующие последовательности:

1) ДПН > ПВИ;
2) ДПН > ПНПР > ПВИ.

Все крысы завершили работу на режиме ПВИ, но у одних была одна история научения (ДПН), а у других — другая (ДПН, а затем ПВИ). Что определит поведение, последний режим (ПВИ) или предшествующая ему история подкрепления? В первом случае между двумя группами не должно возникнуть никаких различий. Ответ представлен на рис. 11.9, на котором сравниваются кумулятивные записи для двух крыс из разных групп. Совершенно ясно, что когда за ДПН сразу следовал ПВИ (рис. 11.9, а), скорость реакции во время ПВИ была низкой. Но использование режима ПНПР заметно влияло на скорость реакции во время ПВИ и приводило к увеличению количества ответов (рис. 11.9, б). Таким образом, на действия при использовании режима ПВИ наиболее сильно влияли события, непосредственно им предшествующие. Режим ДПН влиял на поведение, только если он представлял собой *ближайшее* прошлое (первая крыса). Ученые пришли к выводу, что «полученные данные говорят о преимущественном влиянии непосредственно предшествующего обусловливания на текущее поведение» (р. 549). Такой результат вполне согласуется с представлением о том, что животные быстро адаптируются к изменяющимся окружающим условиям.

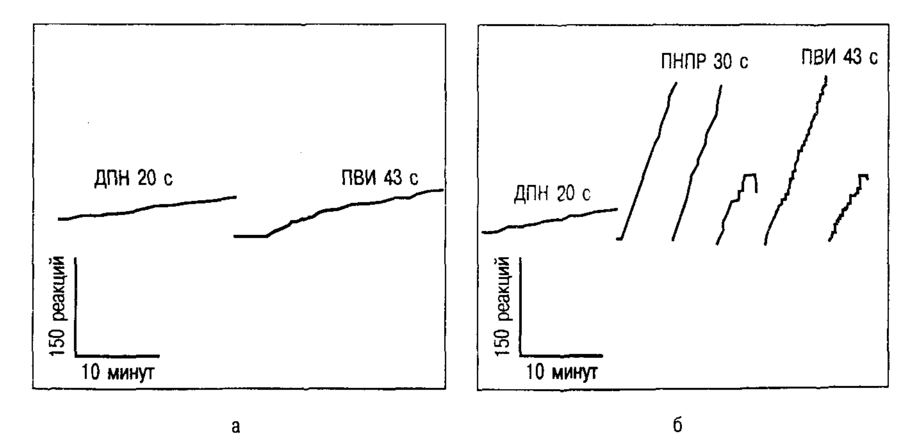


Рис. 11.9. Кумулятивные записи поведения крыс при: а) ДПН > ПВИ или б) ДПН > ПНПР > ПВИ (**источник:** LeFrancois & Metzger, 1993)

Вам может показаться, что этот эксперимент несколько напоминает эксперименты с обобщенными данными, с которыми вы встречались в предыдущих главах. Но хотя в нем использовались две группы по три крысы, общей статистической обработки данных проведено не было и в статье вы не найдете никаких сведений об этом. Вместо этого было описано поведение каждой из шести крыс и представлены шесть кумулятивных записей. В этом эксперименте, конечно, был применен тот же подход, что использовал Дресслер в 90-х гг. XIX в. при изучении лицевого зрения, — явление демонстрируется на одном индивидууме и воспроизводится на других.

Прикладной анализ поведения

По мнению историка Лоренса Смита (Laurence Smith, 1992), есть две категории ученых: одни представляют «созерцательный идеал» и пытаются понять основные причины событий окружающего мира, а другие представляют «технологический идеал» и ищут возможности использовать науку для изменения мира и управления им. Скиннер определенно принадлежал к последней группе. Хотя большинство из его исследований — это чисто лабораторная работа, ученый всегда интересовался применением результатов экспериментального анализа поведения для решения реальных проблем и внес большой вклад в развитие образования и промышленности, воспитание детей и поведенческую терапию. Его идеи использовались даже в космической программе *NASA:* во время двух полетов обезьянам шимпанзе выдавали для изучения сложные оперантные задания. Один из психологов, участовавших в космическом проекте, сказал, что «каждый подход, режим, программирующее или записывающее устройство, которые мы использовали тогда и в дальнейших работах, ведут к Скиннеру или его студентам» (Rohles, 1992).

Кроме того, Скиннер не стесняясь призывал к перестройке общества в соответствии с оперантными принципами, что вызвало огромное число споров и сомнений. Некоторым его советы по улучшению мира казались зловещими, и ученого обвинили в том, что он пытается превратить людей в крыс в проблемных ящиках. Вопрос управления поведением рассматривается во вставке 11.3. Прежде чем переходить к материалу об использовании принципов обусловливания для решения различных прикладных проблем поведения, внимательно прочитайте эту вставку и отметьте для себя, согласны ли вы с предъявленным Скиннеру обвинением.

Прикладную область экспериментального анализа поведения иногда называют **прикладным анализом поведения.** Он включает в себя процедуры, призванные решать жизненные проблемы и основывающиеся на бихевиористских, а в особенности на оперантных принципах. Чтобы дать вам представление о ситуациях, в которых применяются данные принципы, рассмотрим несколько заголовков статей, опубликованных в «Журнале прикладного анализа поведения»:

«Музыкальное подкрепление деятельного поведения па тренировках у пловцов, участвующих в соревнованиях» (Hume & Crossman, 1992);

«Анализ упрощенного лечения заикания у детей» (Wagaman, Miltenberger, & Arndorfer, 1993);

«Исследование отдельных испытуемых при оценке действия напоминаний об использовании ремней безопасности в транспорте: назад к основам» (Berry & Geller, 1991);

«Бихевиористское лечение кофеинизма: снижение чрезмерного употребления кофе» (Foxx & Rubinoff, 1979).

ВСТАВКА 11.3

Этика — управление поведением людей

В главе 1 были описаны цели психологии. Они заключаются в описании, предсказании, объяснении и управлении поведением. Вас может смутить слово «управление», так как оно наводит на мысль о сознательной попытке манипуляции поведением, возможно, даже наперекор желанию человека. Из-за такого возможного подтекста бихевиористов, от Уотсона до Скиннера, обвиняли в том, что они с помощью формирования условных рефлексов хотят полностью управлять поведением людей. Например, в «Записках Клиффа» *[Cliff's Notes),* в которых кратко излагается «Уальден Ту» Скиннера - описание вымышленного общества, основанного на оперантных принципах, - данное общество сравнивается с кошмарным миром романа Оруэлла «1984» (Todd & Morris, 1992). Представление о бихевиористах как о «больших братьях» сильно преувеличено. На его формирование оказывают воздействие, например, случаи, когда читатели встречают название главы из книги Скиннера «Конструирование культуры» или некоторые из утверждений Скиннера вне контекста.

Идея о том, что человек может и должен изменять поведение, исходит из представления бихевиористов о том, что наше поведение во многом обусловлено окружающими условиями. Поскольку окружение *в любом случае* формирует поведение, почему бы не помочь формированию эффективного поведения? Это отношение ясно выражено в двух известных цитатах. Они заключают в себе гораздо более сильную уверенность в решающем влиянии окружающей среды на поведение, чем реально существующую у авторов, и обе показывают, что авторы готовы делать заявления, вызывающие бурные споры. Первое из них - это высказывание Джона Уотсона по поводу воспитания детей: «Дайте мне дюжину здоровых детей и возможность воспитывать их в таких условиях, какие мне потребуются, и я гарантирую вам, что из любого я смогу вырастить специалиста какого угодно профиля - врача, адвоката, художника, продавца, и да, действительно, даже попрошайку или вора...» (Watson, 1924, p. 82).
Вторая цитата из «Уальден Ту» Скиннера. Представляя основателя общества и, вероятно, помня о приведенных выше словах Уотсона, Скиннер написал: «"Что еще нужно сделать?" - сказал он и его глаза заблестели. - "Ну, что вы скажете о конструировании личностей? Это вам интересно? Управление темпераментом? Дайте мне спецификацию м я дам вам человека!"» (Skinner, 1948/1976, p. 274).

Для Скиннера споры по поводу управления поведением не представляли интереса. Для него вопрос о том, следует ли управлять поведением, просто не стоял. Поведение *уже* управляется своими последствиями, и точка. Скиннер считал, что из этого основополагающего факта следует, что необходимо приложить усилия для создания условий, вызывающих эффективное поведение. Критиков это не убедило и они спросили, кто будет принимать решение о том, какое поведение необходимо сформировать. Скиннер решил, что критики потеряли нить рассуждения.

Шумные споры вокруг управления поведением возникли в психологической клинике, где бихевиористские процедуры, с которыми вы скоро познакомитесь, использовались весьма успешно. Одна процедура, вызвавшая наибольшее количество споров, заключалась в применении наказания, в том числе электрического шока, для коррекции поведения детей с серьезными нарушениями психики. Например, в исследовании Кушнера (Kushner, 1970) 7-летнего ребенка с тяжелой умственной отсталостью (его психический возраст составлял 2 года) подвергали электрошоку за кусание рук. Ребенок часто искусывал руки до крови и занес серьезную инфекцию. Попытки сдержать подобное поведение, надевая на него боксерские перчат-

ки или фиксируя положение локтей, не привели к успеху. На бедро ребенку закрепили электроды и подавали электрический разряд каждый раз когда он подносил руку ко рту. Почти сразу наступило изменение поведения, которое стойко держалось даже при удалении электродов.

Означает ли использование подобных процедур в исследованиях по оценке эффективности программ нарушение этического принципа непричинения вреда участникам исследования? Сторонники применения наказания говорят, что другие процедуры зачастую просто не способны изменить такое поведение, как кусание рук и ногтей или удары головий о стены. Если приняты необходимые меры предосторожности (т. е. испробованы и признаны неэффективными другие процедуры, получено обоснованное согласие от родителей или опекунов), суд разрешает использование электрошока «в таких исключительных обстоятельствах, как поведение, направленное на саморазрушение, которое может нанести физический вред субъекту» (Kazdin, 1978, p. 352).

Исследоватэли также утверждают, что для выявления подходящей процедуры необходимо проводить эмпирические исследования, даже если для этого потребуется изменять силу и частоту применяемого наказания. Это значит, что в ходе исследования некоторых детей будут наказывать сильнее, чем предполагается необходимым для изменения их поведения, так как идеальная терапевтическая процедура (т. е. минимальное наказание, требуемое для решения проблемы) может быть определена только после проведения исследования. В свете того, что конечная цель исследований, направленных на оценку использования наказания, заключается в определении наилучшего способа помощи детям с серьезными психическими нарушениями, проведение таких исследований не только оправдано, но и необходимо.

Планы, которые рассматриваются в следующем разделе, чаще всего используются в клинических условиях, но, как показывают приведенные выше названия статей, прикладная работа Скиннера и описанные ниже планы, оиерантные принципы применяются в самых различных обстоятельствах.

Планы с малым *N* в прикладном анализе поведения

В конце отчета об эксперименте по выработке условно-рефлекторного страха у маленького Альберта Уотсоп и Рейнер (Watson & Rayner, 1920) описали несколько способов устранения сформированного страха. Однако ученые не проверяли их на маленьком Альберте. Попытка снизить страх с помощью бихевиористских методов была сделана в ходе новаторского исследования Мэри Джонс (Jones, 1924). У маленького мальчика возрастом 34 месяца по имени Питер Мэри Джонс удалось устранить страх перед кроликами. Она кормила Питера его любимой едой, помещала кролика на некотором расстоянии от него и постепенно сокращала это расстояние. Эта стратегия сходна с применяемыми сегодня процедурами «систематической десенсибилизации».

Бихевиористский подход в терапии стал популярным далеко не сразу после успешного эксперимента Джонс с Питером, но в 50-х, а особенно в 60-х гг. XX в. он получил широкое распространение. Импульсом к распространению этого движения стало подтверждение эффективности процедур, основанных на принципах научения, а также развитие скептического отношения к традиционному подходу в

терапии, особенно к фрейдистским методам (Eysenck, 1952). В 60-х гг. появились журналы, посвященные бихевиористскому подходу в терапии, в том числе «Бихевиористские исследования и терапия» (1963) и «Журнал прикладного анализа поведения» (1968). С этого момента стали регулярно появляться исследования, показывающие, как бихевиористские методы приводят к нужному изменению поведения.

Структура планов с одним субъектом

Структура плана с одним субъектом весьма проста. Поскольку обычно в таких исследованиях отсутствует контрольная группа, необходимо показать, что изменения в поведении отдельного индивидуума наступили в результате применяемого воздействия, а не вследствие влияния каких-либо осложняющих факторов. Это требует выполнения как минимум трех условий. Во-первых, целевое поведение должно быть операционально определено. Недостаточно просто сказать, что производится попытка остановить деструктивное поведение ребенка в классе. Поведение должно быть точно определено в терминах событий, которые легко зафиксировать, например вмешательство в чужую беседу в классе, покидание класса без разрешения и т. д.

Во-вторых, при использовании плана с одним субъектом необходимо установить **базовый уровень** реакции. Это означает, что изучаемое поведение следует наблюдать в течение некоторого времени до воздействия, чтобы определить его обычную частоту (нормальную частоту реакции, по Скшшеру). Именно по базовому уровню реакции можно определить эффективность программы воздействия. Третья часть — это само воздействие и наблюдение поведения. Примите мои поздравления, если вы заметили, что это очень напоминает структуру плана с прерванной временной последовательностью, описанного в предыдущей главе. В обоих случаях цель заключается в том, чтобы оценить эффективность воздействия по отношению к определенному ранее базовому уровню.

Простейший план с одним субъектом иногда называют планом А-В, где А — это базовый уровень, а В — воздействие. Идеальным результатом является изменение поведения при замене А на В. Но я подозреваю, что вы, зная о возможных угрозах внутренней валидности исследования, найдете план А-В довольно слабым. И будете совершенно правы. Изменение поведения может быть вызвано воздействием, но также может быть связано со множеством осложняющих факторов, включая историю, взросление и даже регрессию. Чтобы снизить возможность подобных альтернативных объяснений, используется план с отменой.

Планы с отменой

Если воздействие дает результат и поведение изменяется и при этом изменение связано со взрослением, то маловероятно, что поведение вернется к исходному виду при дальнейшем устранении или отмене воздействия. Но если после отмены воздействия поведение возвращается к базовому уровню, то вероятно, что на него повлияло именно воздействие, а не взросление. Такое рассуждение привело к созданию **плана с отменой** (иногда такой план называется «обратимым»), самым простым вариантом которого является **план А-В-А.** Как вы можете догадаться, этот план начинается аналогично плану А-В, но после того, как воздействие применяется некоторое время, его отменяют (второе А).

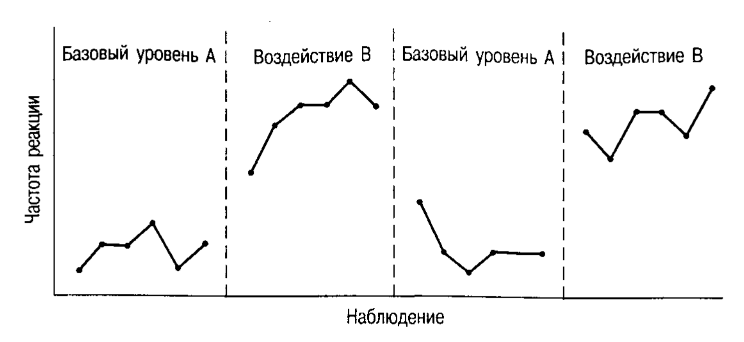


Рис. 11.10. Идеальный результат в случае плана с отменой А-В-А-В

Если изменение поведения строго коррелирует с началом и прекращением воздействия, велика вероятность того, что изменение было вызвано этим воздействием. Уверенность в этом укрепляется, если повторное применение воздействия приводит к аналогичному изменению поведения. Поэтому исследователи предпочитают использовать план А-В-А-В, а не А-В-А. В таком случае программа воздействия проходит по сути двойную оценку. Кроме того, план А-В-А-В имеет этические преимущества, ведь эксперимент завершается активным воздействием на участников. Идеальный результат такого эксперимента представлен на рис. 11.10. Обратите внимание, что для того, чтобы воздействие было признано эффективным, после его отмены поведение должно вернуться к базовому уровню (или близко к нему), а при возобновлении воздействия поведение должно вновь измениться. Подобные результаты трудно объяснить иначе чем успешным применением программы воздействия. Для иллюстрации использования этого весьма распространенного плана рассмотрим пример с пловцами.

Пример 28. План А-В-А-В

Исследование Хьюма и Кроссмана (Hume & Crossman, 1992) направлено на совершенствование продуктивной манеры поведения у пяти пловцов в возрасте от 12 до 16 лет в ходе «сухопутного тренировочного периода» (р. 665) — когда молодые люди не тренировались в воде, но должны были выполнять определенные упражнения для развития навыков плавания. «Продуктивное» поведение включало упражнения (например, подъем туловища из положения лежа) или показ различных техник другим пловцам. «Непродуктивное» поведение включало неуместные упражнения (например, стояние на руках), отвлекающую активность (например, попытки спрятать чужие очки) и обсуждение не относящихся к делу вопросов (например, разговоры о девушках). Как вы можете догадаться, особенно принимая во внимание возраст молодых людей, базовый уровень непродуктивного поведения был довольно высок. Это вызывало некоторое раздражение у тренеров, которые, по-видимому, забыли, что когда-то и им было по 12-16 лет.

Три обученных наблюдателя (почему три?[1]) фиксировали поведение во время базового периода (А), состоящего примерно из десяти тренировок. Затем програм-

Чтобы ответить на этот вопрос, загляните в главу 12.

ма подкрепления (В) была введена на несколько занятий, отменена (А) еще на несколько тренировок и снова введена (В). Трое молодых люден были распределены в группу «ситуативного подкрепления», остальные двое вошли в группу «неситуативного подкрепления». Подкреплением было проигрывание популярной среди мальчиков (это выяснили заранее) музыки во время тренировки. Чтобы музыка играла (для всех пяти мальчиков) во время следующей тренировки, трое мальчиков из группы условного подкрепления должны были достигнуть определенного уровня продуктивного поведения во время занятия вне бассейна.

Как это обычно делается в оперантном направлении исследований для планов с малым *N,* результаты были представлены по каждому пловцу в отдельности. На рис. 11.11 показаны результаты одного пловца из группы условного подкрепления — они говорят о том, что для него программа оказалась довольно эффективной. В базовый период непродуктивное поведение преобладало на всех тренировках, кроме одной, но после введения программы это быстро изменилось и продуктивное поведение стало появляться чаще, чем непродуктивное. Действие отмены программы и ее последующего возобновления подтвердило, что музыкальное подкрепление поведения повышает качество тренировок.



Рис. 11.11. Данные исследования Хьюма и Кроссмана (1992), полученные при использовании плана А-В-А-В для увеличения продуктивности действий пловцов

Планы с несколькими базовыми уровнями

Иногда планы с отменой не подходят для проведения исследования. Например, если программа включает формирование определенного навыка, этот навык остается и после прекращения программы. Другими словами, при отмене воздействия поведение не возвращается к базовому уровню, а держится на достигнутом уровне. План с отменой также может вызывать этические проблемы, особенно если изменяемое поведение направлено на саморазрушение. Если группа воздействия состоит из детей с серьезными нарушениями психики, которые постоянно бьются головами о стены, и с помощью оперантной процедуры удается остановить подобное поведение, то отмена воздействия для проверки, возобновится ли данное поведение, будет не вполне оправданной.

В случае **плана с несколькими базовыми уровнями** устанавливаются базовые уровни, а затем *в различное время* применяется воздействие. Есть три возможности — базовые уровни могут быть установлены: а) для одного типа поведения у двух

или нескольких индивидуумов, б) для двух или нескольких типов поведения у одного индивидуума и в) для одного типа поведения у одного индивидуума, но в различных условиях. Структура плана одинакова во всех трех случаях. Рассмотрим, к примеру, ситуацию, в которой требуется изменить три типа поведения у одного индивидуума. Идеальный результат представлен на рис. 11.12, позаимствованном из исследования Рени и Крира (Renne & Creer, 1976).

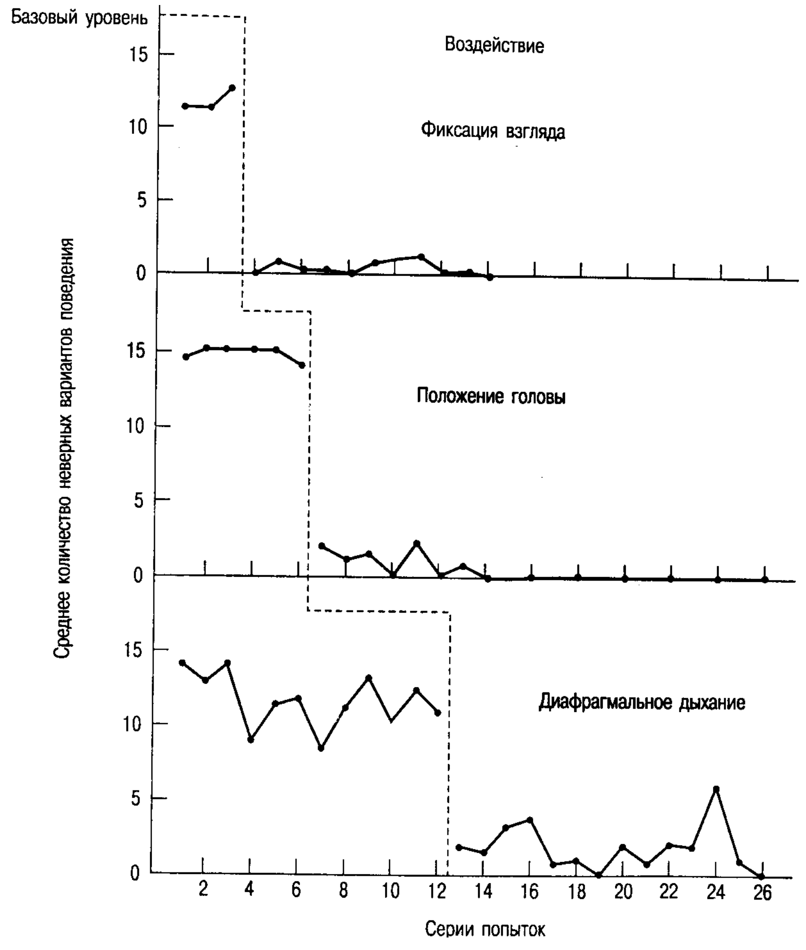


Рис. 11.12. План с несколькими базовыми уровнями, использованный для изменения трех типов поведения у одного индивидуума

Авторы предприняли попытку научить страдающих астмой детей пользоваться дыхательным аппаратом и выявили три типа «неверного» поведения, которые следовало устранить, чтобы дети начали правильно пользоваться устройством. Как видно по рисунку, исследование началось с определения стандартного базового уровня для каждого из трех типов поведения. Затем для одного из них было введено воздействие. Если воздействие эффективно, то оно должно было повлиять толь-

ко на первый тип поведения. Аналогично, когда воздействие распространилось на поведение 2, оно должно было повлиять только на этот тип поведения, но не на поведение 3. В момент времени 3 должен был измениться третий вид поведения. Таким образом, вывод об эффективности программы делался на том основании, что она вводилась в разное время, а изменения в поведении появлялись на конкретном отрезке времени и не раньше. Если бы все три или два из трех типов поведения изменялись после начала воздействия в момент времени 1, то эти изменения нельзя было бы связать с применением воздействия. В исследовании Рени и Крира поведением 1 была фиксация взгляда — он должен был сфокусироваться на ручке манометра, причем голова должна была быть повернута под углом 90° к аппарату. Любое другое положение головы, кроме угла в 90°, считалось неверным. Второй тип поведения заключался в отсутствии раздувания ноздрей, свидетельствующего о дыхании через нос, а не через рот. Поведение 3 состояло в использовании особого вида дыхания. Правильное поведение подкреплялось с помощью системы оценок, соответствующих различным вознаграждениям. Как показано на рис. 11.12, на изменение условий дети мгновенно реагировали фиксацией взгляда и изменением положения головы. Дыхание с помощью диафрагмы также возникало, хотя в этом отношении поведение изменялось не до конца (это было самое трудное из трех заданий).

В литературе по прикладному анализу поведения исследования с несколькими базовыми уровнями встречаются очень часто. Хороший пример тому — исследование по оценке программы помощи заикающимся детям (Wagaman, Miltenberger & Arndorfer, 1993), в котором использовался план с несколькими базовыми уровнями. Рассмотрим это исследование более подробно.

Пример 29. Несколько базовых уровней у нескольких испытуемых

Данное исследование иллюстрирует первый из описанных выше видов плана с несколькими базовыми уровнями. Оно было направлено на изменение одного и того же типа поведения у нескольких индивидуумов. Стратегия исследования заключалась в изучении влияния на индивидуума определенной программы и в повторении ее с другими участниками. Исследуемым поведением было заикание, испытуемыми — восемь школьников (шесть мальчиков и две девочки) в возрасте от 6 до 10 лет. Программа воздействия включала упрощенный вариант процедуры «регулируемое дыхание». В ее ходе детей обучали координировать дыхание и речь. Обучение проходило на дому; родителей также учили пользоваться данной техникой. Кроме того, программа включала элементы социального подкрепления (со стороны родителей) и обучающий тренинг, в ходе которого родителей и детей учили быстро идентифицировать различные виды речевых проблем.

Использование подобного тренинга подразумевает определение заикания. Очевидно, что программа, ставящая целью полное осознание испытуемыми заикания, Должна начинаться с выработки четкого операционального определения данного явления. Уэгамэн и др. (1993) разработали следующую классификацию видов заикания: «а) повторение слов, б) повторение частей слов, в) пролонгация звука и г) остановка или прерывистость речи до завершения слова» (р. 55). Для повышения надежности измерений исследователи провели аудиозаписи всех занятий и предоставили записи для оценки нескольким компетентным лицам. Уровень согласия оценивающих был высоким.

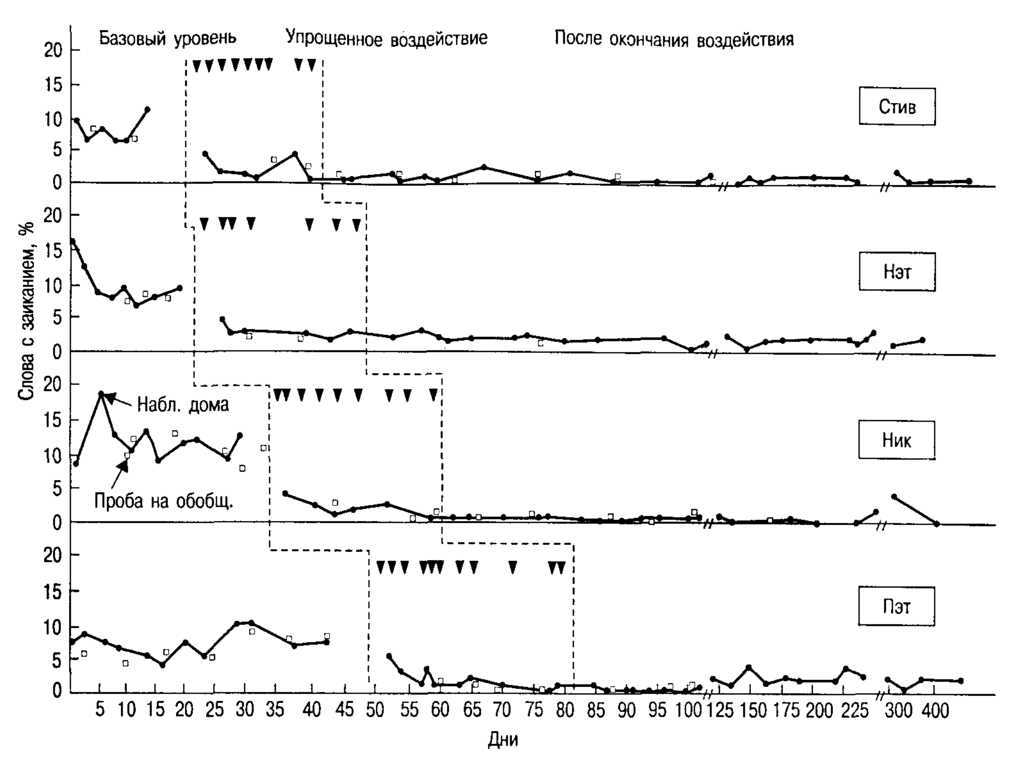


Рис. 11.13. Снижение заикания в исследовании с несколькими базовыми уровнями, проведенном Уэгамэном, Милтенбергером и Арндорфером (1933)

На рис. 11.13 показаны результаты четырех детей (для остальных четырех участников были получены аналогичные кривые). Следует отметить несколько моментов. Во-первых, как это делается в случае планов с одним субъектом, исследование началось с установки базового уровня, в данном случае нескольких базовых уровней. Во-вторых, и это характерно для планов с несколькими базовыми уровнями, программа запускалась в различное время. Обычно представленный результат исследования с несколькими базовыми уровнями можно опознать по нескольким графикам, расположенным один над другим, и пунктирной линии, идущей ступенями от одной кривой к другой. В-третьих, воздействие явным образом сказывалось на заикании каждого ребенка.

Если вы сосчитаете маленькие стрелочки, расположенные над участками, соответствующими моментам действия программы, то обнаружите, что их количество различается. У Стива, к примеру, было девять занятий по программе, а у Нэт — только семь (имена, конечно, изменены). Поскольку вы изучили предыдущие главы, вы, вероятно, заинтересуетесь вопросом контроля над данным исследованием. Нужно ли, чтобы количество занятий было одинаковым у всех участников? Не забывайте, что основа планов с малым Л' — это изучение отдельных индивидуумов. Уэгамэн и др. (1993) больше заботились о том, чтобы каждый ребенок достиг определенного уровня поведения, чем о поддержании постоянным общего количества занятий. Поэтому

занятия с каждым ребенком продолжались до тех пор, пока каждый из них «не достиг критериального уровня заикания (<3% слов с заиканием)» (р. 57).

Процедуры с малым *N* иногда критикуют за то, что их результаты нельзя распространить за пределы экспериментальной ситуации, и за отсутствие дополнительных исследований. В исследовании Уэгамэна и др. обе эти проблемы были решены. Во-первых, обратите внимание, что на рис. 11.13 большинство точек обозначено кругами, а некоторые — квадратами. Последние помечены как «проба на обобщ.», что значит «проба на обобщение». Они указывают моменты, когда исследователи оценивали заикание детей в школьных условиях. Круги («набл. дома») обозначают занятия дома. Очевидно, что хотя занятия проходили дома, их результаты можно распространить и на условия школы. Во-вторых, отметьте, что, как показывает на рис. 11.13 участок «после окончания воздействия», для каждого ребенка было проведено большое дополнительное исследование (10-13 месяцев). Таким образом, эффект программы был не краткосрочным, а длительным.

Планы с изменяющимся критерием

Третий из важнейших видов планов с одним субъектом — **план с изменяющимся критерием** (Hartmann & Hall, 1976), в основе которого лежит оперантная процедура, носящая название формирование. Формирование поведения с помощью этой процедуры предполагает подкрепление последовательных приближений к конечному желаемому поведению. В таком плане целевое поведение является слишком сложным, чтобы испытуемый мог достичь его сразу, и поэтому его приходится формировать постепенно. Процедура начинается с установки базового уровня; далее вступает в силу воздействие, которое продолжается до тех пор, пока не будет достигнут некий начальный критерий. Затем критерий становится все более точным, пока не будет сформировано целевое поведение. Связанные со сферой здоровья программы, например диетические или программы физических упражнений, — это отличные кандидаты для использования таких планов. К примеру, Фокс и Рубинофф (Foxx & Rubinoff, 1979) в своем исследовании попытались решить хорошо известную проблему чрезмерного потребления кофеина. Для тех, кто выпивает в день по 15-20 чашек кофе, сразу перейти на 2-3 чашки просто невозможно. Гораздо лучше снижать количество чашек шаг за шагом и использовать особые вознаграждения за достижение новых все более сложных критериев. Эту стратегию успешно применили Фокс и Рубинофф. В примере 30 план с изменяющимся критерием был использован для решения другой распространенной проблемы. Исследование состояло в улучшении физического состояния детей с лишним весом.

Пример 30. План с изменяющимся критерием

Не секрет, что борьба с лишним весом — это одна из навязчивых идей американцев. В данном случае она не лишена оснований. Особенно удручает количество детей, имеющих данную проблему, — по некоторым оценкам, один из каждых четырех детей в Америке страдает ожирением (Gleitman, Fridlund & Riesberg, 1999)[1]. Ситуация часто заметно осложнена отсутствием физической активности. Делюка и Холборн в своем исследовании (DeLuca & Holborn, 1992) использовали план с изменяющимся крите-

[1] Страдающими ожирением считаются люди, чей вес превышает средний для их роста на 20% или более.

рием и элементы плана с отменой. Они изучили возможность формирования физической активности у трех страдающих ожирением и трех не имеющих этой проблемы 11-летних мальчиков. Упражнения выполнялись на велотренажере, который звонил и мигал лампочками, когда участник зарабатывал подкрепление. Исследование началось с установки базового уровня. В ходе восьми последовательных попыток дети должны были «заниматься столько, сколько захотят» (р. 672). После того как средний базовый уровень был установлен и определен как среднее число вращений педалей в минуту, был назначен первый критерий — на 15% выше базового уровня. Обратите внимание, что в соответствии с методологическими принципами малого *N*, заключающимися в концентрации на отдельном индивидууме, первый критерий (а также и все последующие) не был одинаковым для всех детей, а определялся в зависимости от начального значения, полученного каждым испытуемым.

После того как первый критерий был установлен, приступили к занятиям по режиму с изменяемым подкреплением. Мальчиков снова попросили заниматься столько, сколько они пожелают, но на этот раз, если они крутили педали со скоростью, на 15% превышающей базовый уровень, раздавался звонок и загоралась лампочка. Это приносило очки, набрав определенное количество которых дети могли получить ценные призы (например, комиксы). Через восемь занятий критерий увеличился еще на 15%, а затем увеличился еще раз. За этим последовали три занятия в режиме отмены, во время которых подкрепление было временно прекращено. Исследование завершилось возвращением к критериальному уровню, достигнутому до отмены подкрепления. На рис. 11.14 показаны результаты, полученные всеми шестью мальчиками. Как видно, уровень выполнения упражнений и у одних, и у других мальчиков постепенно увеличился, за исключением Перри, который пропустил несколько занятий (отмеченных пунктирной линией). Также видно, что уровень выполнения упражнений снизился при отмене подкрепления (этап отмены на графике обозначен БУ, что значит «базовый уровень»). Обратите внимание на слабое место данной работы — отсутствие каких-либо дальнейших исследований. Как вы, возможно, знаете из личного опыта, программы физических упражнений очень трудно реализовывать в течение долгого времени. Интересно было бы посмотреть, отличается ли продолжительность эффекта в случае такого оперантного подхода.

Данное исследование иллюстрирует еще две особенности прикладного анализа поведения. Во-первых, в исследовании был решен вопрос о том, каким должно быть подкрепление, ведь некоторые мальчики могут захотеть работать за сборник комиксов, а другие — нет. Чтобы дети выполняли задание, получая одинаково ценное подкрепление, Делюка и Холборн попросили их в начале исследования заполнить «анкету подкрепления», а затем по десятибалльной шкале оценили, как сильно мальчикам нравятся определенные вещи. Затем каждый мальчик работал с таким подкреплением, которое было им высоко оценено. Во-вторых, исследователи решили проблему **социальной валидности** (Wolf, 1978) (термин из прикладного анализа поведения). Этот вид валидности говорит: а) имеет ли конкретный анализ поведения ценность с точки зрения совершенствования общества, б) осознается ли его ценность участниками исследования и в) в какой степени участники действительно пользуются возможностями программы (Geller, 1991). Делюка и Холборн оценили социальную валидность, попросив каждого из мальчиков, их родителей и

учителей заполнить «опросник социальной валидности» (р. 673). Полученные результаты они охарактеризовали как «равномерно положительные». Другие показатели успешности программы были единичными: все мальчики впоследствии катались на велосипедах, в ходе программы все страдающие ожирением дети уговорили своих родителей купить им новые велосипеды и все участники расстроились, когда программа завершилась. Как было отмечено ранее, результаты можно подкрепить, проведя дополнительное исследование примерно через 6 месяцев.

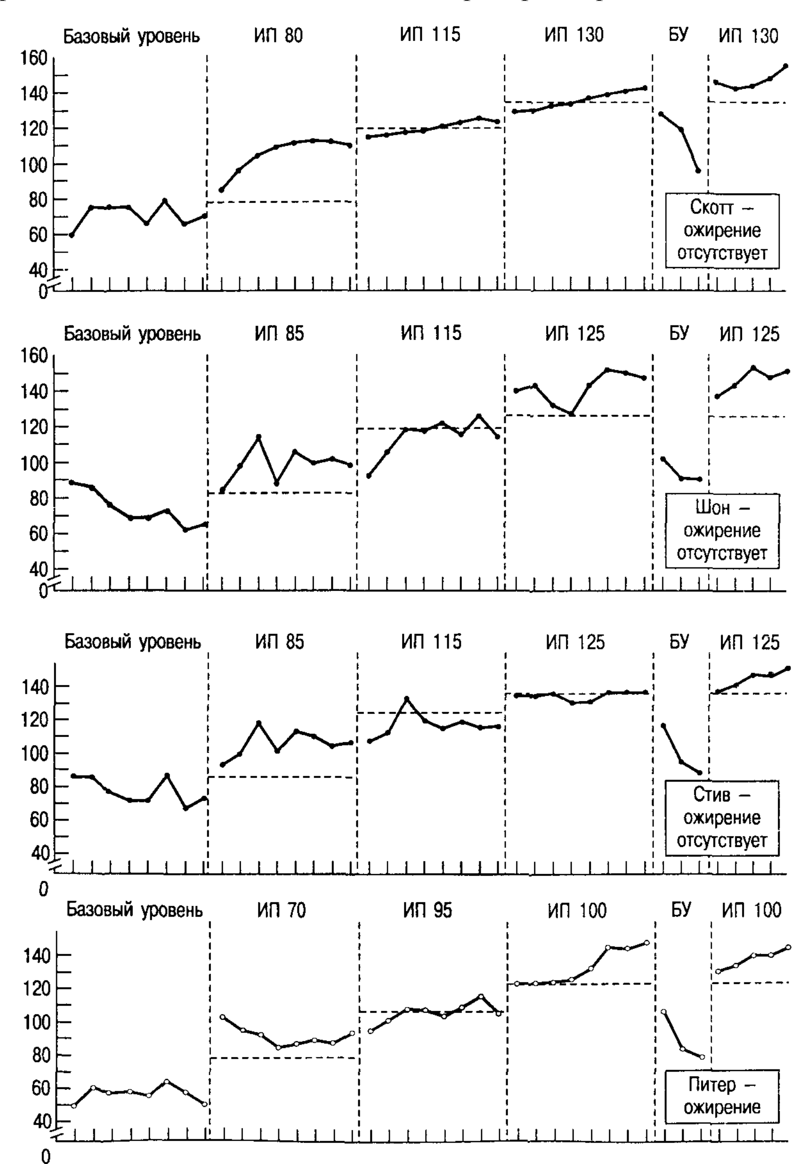


Рис. 11.14. Данные направленного на улучшение физического состояния 11 -летних мальчиков с ожирением и без него исследования по плану с изменяющимся критерием (из DeLuca & Holborn, 1992)

Другие планы

В прикладном анализе поведения используются не только планы с отменой А-В-А-В, планы с несколькими базовыми уровнями или изменяющимся критерием. В зависимости от изучаемой проблемы, многие исследователи объединяют элементы основных планов, несколько изменяют их или создают новые. Один из модифицированных планов А-В-А-В — **план А-В-С-В.** Его часто используют в ситуациях, когда в ходе программы используется условное подкрепление. Во время воздействия (В) целевое поведение подкрепляется сразу после своего появления, а в другое время подкрепление не используется. В период С подкрепление осуществляется, но оно не зависит от наличия целевого поведения. Таким образом, подкрепление осуществляется как при В, так и при С, но зависит оно только от поведения, демонстрируемого на этапе В. Такой план позволяет выявить эффект плацебо, показывая, что поведение изменяется не просто потому, что субъект получает удовольствие от подкрепления, доволен, что ему уделяется столько внимания, и т. д., но также из-за наличия определенных условий подкрепления.

Другой пример модифицированного плана А-В-А-В — **план А-А^В-Аj-В.** Его часто используют при оценке эффективности медикаментозного лечения в случае одного индивидуума. Используются и другие последовательности, но эта является наиболее распространенной (например, Liberman, Davis, Moon, & Moore, 1973). А — это обычный базовый уровень, а А, — второй базовый уровень, когда участники получают плацебо. Настоящее лекарство выдается на этапе В.

Еще один пример плана с одним субъектом — план с чередующимся воздействием. Его используют при сравнении влияний более чем одного вида лечения на одного индивидуума. Этот тип плана стал популярен, поскольку он позволяет в ходе одного исследования оценить более чем один вид воздействия. После установления обычного базового уровня попеременно применяются различные стратегии воздействия, причем они чередуются несколько раз и располагаются в случайном порядке, чтобы снизить возможный эффект последовательности (как вы догадываетесь, это еще один вариант позиционного уравнивания).

Оценка планов с одним субъектом

Рассмотренные нами планы очень помогают при оценке эффективности оперантного и других подходов к условно-рефлекторному изменению поведения. Все они берут начало от утверждения Павлова/Скиннера, что при строгом контроле условий возникает систематическое и предсказуемое поведение. Действенность этого подхода была продемонстрирована в различных ситуациях, начиная от терапевтического изменения поведения отдельных индивидуумов до изменений в поведении общества. Однако бихевиористские планы с малым N *не* лишены недостатков.

Чаще всего критика касается вопроса внешней валидности, или возможности обобщения результатов. Если обнаружено, что конкретный вид терапии поведения эффективен для одного человека в конкретной ситуации, то как узнать, эффективен ли он для других людей, имеющих ту же проблему? Возможно, участвовавший в исследовании человек имел какие-то особенности. Или, может быть, эффект от воздействия, появившийся в одной ситуации, нельзя обобщить на другие условия.

Сторонники изучаемого нами подхода говорят, что в действительности в некоторых исследованиях возможность обобщения оценивается. Хорошим примером тому является исследование Уэгамэна и др. (Wagaman et al, 1993) по преодолению заикания у детей, в котором использовался план с несколькими базовыми уровнями. Во-вторых, необходимые повторение и дополнение в действительности весьма распространены в исследованиях с одним субъектом. Например, техника формирования поведения с помощью «дифференциального внимания» (родители уделяют внимание желаемому поведению детей и игнорируют нежелательное) сегодня хорошо разработана благодаря проведению десятков исследований с малым *N,* показавших свою эффективность для различных видов поведения. Рассмотрев только данные по популяции маленьких детей, Барлоу и Херсен (Barlow & Hersen, 1984) получили список из 65 исследований, посвященных дифференцированному вниманию, опубликованных с 1959 по 1978 г.

Кроме того, планы с одним субъектом часто критикуют за отсутствие статистического анализа и оценку результатов исследования на основании простого обзора данных. В какой-то степени это отражает разницу методологических принципов сторонников исследований с большим и малым *N*. Защитники планов с малым Договорят, что выводы делаются только тогда, когда эффект достаточно силен, чтобы быть очевидным для всех. Однако стоит отметить, что в исследованиях с одним субъектом стали появляться элементы статистического анализа. Например, из исследований по оценке программ пришел анализ временных рядов, который используется в некоторых исследованиях с одним субъектом для отделения результатов воздействия от результатов влияния тенденций (Junginger & Head, 1991). Также анализ временных рядов помогает решить проблему относительно нестабильных базовых уровней, которая затрудняет визуальное изучение графиков в исследованиях с одним субъектом. Статистический анализ даже смог проскользнуть в «Журнал экспериментального анализа поведения» — самый скиннерианский из всех журналов. Один исследователь оперантного направления сокрушался, что при изучении статей, опубликованных в этом журнале в 1989 г., он обнаружил, что около трети всех статей включают тот или иной вид статистического анализа и не более 10% статей содержат кумулятивные записи (Baron, 1990).

В-третьих, планы с одним субъектом критикуют за то, что исследователи не проводят достаточной проверки на эффект взаимодействия. Как вы помните из главы 8, одна из самых привлекательных особенностей факторных планов состоит в возможности выявления взаимодействий между двумя или более независимыми переменными. Планы взаимодействия для малого *N* существуют, но они очень громоздки. Например, в исследовании Лайтенберга и др. (Leitenberg et al., 1968, цитируется в Barlow & Hersen, 1984) для сравнения двух методов терапии (В и С) и оценки их совместного влияния был использован план A-B-BC-B-A-B-BC-B. Но обратите внимание, что метод С сам по себе не встречается в плане. Поэтому требуется повторение исследования по плану A-C-BC-C-A-C-BC-C с участием другого субъекта.

Один особо интересный вид взаимодействия, с которым вы познакомились в главе 8, нередко появляется в плане Р χ Е, содержащем как субъектную (Р), так и управляемую (Е) переменные. Один из вариантов взаимодействия Р χ Е возника-

ет, если управляемый фактор влияет на один тип людей одним образом, а на другой — другим. Субъективные переменные в плане Р χ Е, конечно, межсубъектные, но, за исключением некоторых исследований с несколькими базовыми уровнями, планы с одним субъектом являются внутрисубъектными. Таким образом, взаимодействия Р χ Е, аналогичные только что описанным, можно найти только в односубъектных планах со сложными многократными повторениями, в которых обнаруживается, что: а) воздействие 1 эффективно в случае испытуемого 1-го типа, но не в случае испытуемого 2-го типа и б) воздействие 2 эффективно в случае испытуемого 2-го типа, но не в случае испытуемого 1-го типа.

Также исследования с одним субъектом, проводимые последователями оперантного направления, подвергаются критике за то, что все внимание в них уделяется частоте реакции, поскольку именно она является зависимой переменной. Данное направление не включает исследования, в которых изучается скорость реакции, правильность запоминания слов, время, потраченное на наблюдение (как при исследовании привыкания), или другие зависимые переменные, проливающие свет на законы поведения. Частота реакции — это, несомненно, очень важная переменная, но при этом нельзя недооценивать значение других параметров.

Планы с малым *N* встречаются не только в оперантных исследованиях и прикладном анализе поведения. Один из самых старых психологических подходов — изучение порогов чувствительности с помощью методов психофизики.

Психофизика

В начале этой книги я указал на учебник Вильгельма Вундта, выпущенный в 1874 г., как на поворотный пункт в становлении психологии как научной дисциплины. Но экспериментальная психология существовала и до Вундта, чьей научной деятельности предшествовала работа Густава Фехнера (1801-1887). Его книга «Элементы психофизики» считается первым учебником по экспериментальной психологии. Его исследования положили начало развитию **психофизики** — изучению взаимосвязей между физическими стимулами и их ощущением.

Подробное описание исследований по психофизике не входит в задачи данной главы. Вы можете получить более, полные сведения об этом, пройдя курс по ощущениям и восприятию. Наша задача заключается в том, чтобы сделать краткий обзор видов психофизических исследований и показать еще одну возможность использования планов с малым *N*.

Пороги

В исследованиях по психофизике рассматриваются два взаимосвязанных вида навыков: способность к обнаружению стимулов и способность к дифференциации едва различающихся стимулов. Проблема обнаружения стимула изначально связана с определением **абсолютного порога** (термин, введенный Фехнером) — интенсивности стимула, достаточной для его обнаружения. Каждый раз, когда вы проверяете остроту зрения или слуха, вы участвуете в эксперименте по изучению абсолютного порога. Поскольку известно, что абсолютные пороги очень редко или вовсе не являются пределом восприятия, современные исследования чаще всего

проводятся в соответствии с так называемой «теорией обнаружения сигналов» и направлены на выявление: а) факторов, влияющих на способность обнаружения стимулов, и б) особенностей влияния таких индивидуальных переменных, как мотивы, внимание и усталость на процесс принятия решения о наличии стимула.

Второй вид порога, который изучается в психофизике, связан с различением и называется **дифференциальным порогом,** или «едва заметным различием» (jnd от *англ. just noticeable difference).* Если вы смотрите на две лампочки, одна из которых немного ярче другой, но не замечаете разницы, то это значит, что дифференциальный порог еще не был достигнут. Если интенсивность одной из лампочек увеличить так, чтобы вы заметили разницу, то будет достигнут дифференциальный порог. С приобретением опыта дифференциальный порог может измениться, что даст возможность проводить более тонкое различение стимулов. Это становится очевидным при сравнении профессиональных настройщиков фортепиано или дегустаторов вина (а в XIX в. сортировщиков шерсти) с обычными людьми.

Методы психофизики

Фешнер описал три основных психофизических метода, которые используются и сегодня. Представьте, что вы проходите проверку слуха. Вы сидите в наушниках в звуконепроницаемой кабине и каждый раз, когда слышите звук, поднимаете руку. Исследователь подает сигнал, который хорошо слышен, затем более слабый, потом еще слабее, пока вы не перестаете его слышать. Через некоторое время вам кажется, что вы опять слышите звук, далее следует еще один, немного более громкий, потом еще громче и т. д. В этом случае исследователь пользуется **методом минимальных изменений (ММИ),** который состоит из чередующихся «нисходящих» и «восходящих» попыток. При нисходящих попытках первый стимул заметно превышает порог, а интенсивность последующих стимулов постепенно снижается, пока они не перестанут обнаруживаться. При восходящих попытках интенсивность первого стимула ниже пороговой, а далее постепенно возрастает. Общий порог определяется нахождением среднего между двумя значениями интенсивности: одним, при котором звук начинает различаться при восходящих попытках, и другим, при котором звук перестает различаться при нисходящих.

Два других подхода для определения порогов — это метод постоянных стимулов и метод подгонки. При изучении слуха **методом констант** интенсивность сигнала при конкретной попытке определяется случайным образом. Это позволяет избежать предположений и догадок, которые могут возникнуть при использовании метода пределов. **В методе средней ошибки** субъект непосредственно управляет интенсивностью стимулов и регулирует ее до тех пор, пока стимул не становится различимым, т. е. не начинает отличаться от второго стимула.

В заключение отметим, что оперантные методы также используются для изучения порогов восприятия у животных. Такие исследования проводятся, к примеру, с использованием процедур дифференциации. Если животное способно нажимать рычаг при звуке 1 и не нажимать его при звуке 2, то оно может различать эти два сигнала. Если оно не может научиться дифференцировать звуки, то, вероятно, оно не слышит разницы между ними.

Психофизика и малое *N*

Немного поразмыслив о перечисленных методах, вы поймете, почему исследования по психофизике почти всегда проводятся с малым числом участников. Во-первых, для проведения многих психофизических исследований требуется участие очень небольшого количества испытуемых, так как сенсорные системы большинства людей очень похожи. Поскольку индивидуальные различия незначительны, результаты, полученные для нескольких участников, можно легко обобщить на остальных людей. И конечно, вынесенный вердикт, как обычно, подтверждается повторением.

Вторая причина использования в психофизике планов с малым N связана с различием внутрисубъектных и межсубъектных факторов. Задача по установлению порогов требует использования внутрисубъектных планов, в которых каждый испытуемый подвергается воздействию большого количества стимулов разной интенсивности. Нет никакого смысла предлагать одному испытуемому стимул 1, другому — стимул 2 и т. д., а затем вычислять средний порог для большой группы людей. Пороги необходимо определять для отдельных людей, а это можно сделать, только проведя большое количество измерений с широким диапазоном стимулов для одного индивидуума. Кроме того, бессмысленно тратить время участников и предлагать им по одному или два стимула, ведь одна попытка занимает лишь несколько секунд. Кто захочет проводить исследование, в котором требуется 5 минут для настройки оборудования, 5 минут для инструктажа участников, 10 минут на дебрифинг и 30 *секунд* на сбор данных?

Третья причина проведения психофизических исследований с малым JV связана с контролем. Большинство исследований порогов проводятся в строго контролируемых лабораторных условиях. Так при каждой попытке полученные данные будут обладать меньшей дисперсией ошибки, чем в случае, скажем, эксперимента по социальной психологии, исследующего готовность к оказанию помощи. Это, конечно, снова напоминает нам о Скиннере: при наличии контроля за условиями возникает точный результат.

В этой главе вы познакомились с психологическим направлением, которое имеет богатую историю и решительных сторонников. Данный подход идеален для проведения исследований определенного типа: например, для изучения влияния подкрепления на обучение или определения сенсорных порогов. В других обстоятельствах, когда необходимо использовать межсубъектный план или требуется проводить большое количество измерений зависимой переменной, этот подход менее пригоден. В следующей главе, завершающей рассмотрение методов исследований, изучаются описательные подходы.

Резюме

Первые исследования по психологии и малое *N*

До появления сложных статистических методов психологические исследования проводились в основном с участием отдельных индивидуумов или небольшого количества испытуемых, а экспериментатором обычно являлся один из участни-

ков. В исследованиях, включающих более одного участника, все испытуемые, кроме первого, использовались с целью повторения исследования. Полученные данные представлялись по отдельности и редко обобщались для получения средней оценки для группы людей.

Причины использования планов с малым *N*

Сегодняшние защитники исследований с использованием одного или малого числа участников утверждают, что обобщенные данные могут неверно отображать индивидуальные действия или особенности и вводить исследователя в заблуждение относительно изучаемого явления. Во-вторых, при изучении определенных явлений (например, редких психических нарушений) бывает невозможно набрать большое количество участников. Третий аргумент основывается на мнении Скиннера: для лучшего понимания, предсказания и управления поведением необходимо подробно изучать отдельных индивидуумов. Скиннер считал, что если условия поведения контролируются в достаточной мере, то оно становится предсказуемым и появляется возможность измерения.

Планы с малым *N* в прикладном анализе поведения

Прикладным анализом поведения называется применение принципов бихевиоризма для изменения поведения людей, которое нередко проводится в клинических условиях. В таком случае используются планы с одним субъектом. Наиболее часто используется план с отменой, в котором до начала программы измеряется базовый уровень поведения. Через некоторое время воздействие прекращается (отменяется), а затем вновь восстанавливается (план A-B-A-B). Вывод об эффективности воздействия делается, если отмена и восстановление воздействия непосредственно влияют на поведение, которое требуется изменить. Другие варианты плана с одним субъектом — это план с несколькими базовыми уровнями, который используется, если отмена невозможна, и план с изменяющимся критерием, при котором целевое поведение, которого нельзя достичь единовременно, формируется постепенно.

Психофизика

Кроме исследований, проводящихся в соответствии с оперантным подходом, есть и другие разновидности исследований с малым *N*. Лучший пример тому — психофизика, занимающаяся изучением взаимосвязей между физическими стимулами и их обнаружением. Традиционные исследования по психофизике направлены на определение абсолютных и дифференциальных порогов восприятия; такую работу ученые проводили уже на заре развития экспериментальной психологии. Фешнер первым выделил три метода, применяющихся и в настоящее время: ММИ, метод констант и метод средней ошибки. Современная психофизика под влиянием теории обнаружения сигналов занимается изучением порогов как с точки зрения способностей наблюдателя к ощущению стимулов, так и со стороны неперцептивных факторов, таких как критерии принятия решений.

Задания для повторения

Выбор ответа

1. Какое из следующих утверждений об исследовании лицевого зрения Дресслера верно?
   1) оно показывает, что обобщение данных нескольких испытуемых может дать результат, который не будет отражать поведение отдельных испытуемых;
   2) оно является одним из первых примером плана с малым *N,* в котором дополнительные участники изучаются с целью повторения исследования;
   3) оно является примером психофизического метода пределов;
   4) оно представляет собой исключение из того правила, что на заре психологии использовались планы с малым *N.*

2. Все нижеприведенные аргументы обычно приводятся защитниками использования планов с малым *N, кроме следующих:*
   1) вследствие индивидуальных различий с помощью статистических методов необходимо определить, является ли независимая переменная причиной изменений зависимой переменной;
   2) иногда, при осреднении данных нескольких участников, полученная картина не отражает поведения отдельных индивидуумов;
   3) исследования сравнительно редких психологических явлений бывает возможно провести только при использовании плана с малым *N;*
   4) если строгий контроль за условиями дает упорядоченное поведение, то особое внимание следует уделить контролю за окружающими условиями для одного организма, а не осуществлению статистического контроля.

3. В каких случаях план с несколькими базовыми уровнями имеет преимущество перед планом A-B-A-B?
   1) если целевое поведение нельзя получить сразу;
   2) если цель заключается в сравнении влияния двух различных стратегий воздействия на одного участника;
   3) если по какой-либо причине невозможна отмена воздействия;
   4) если необходимо исследовать более одного индивидуума.

4. Исследователь хочет разработать и провести оценку программы, в ходе которой дети из учреждения для умственно отсталых должны научиться правильно заправлять по утрам кровати. Персонал учреждения рекомендует два подхода. Один включает использование системы начисления очков, когда дети зарабатывают очки и получают за это награды. Второй подход состоит в использовании в качестве подкрепления внимательного отношения персонала. В таком случае если ребенок правильно заправляет кровать, то ему

уделяют больше внимания. Какой план лучше использовать, чтобы оценить относительную эффективность этих подходов?

1) план с чередующимся воздействием;
2) план с отменой;
3) план с несколькими базовыми уровнями;
4) план с изменяющимся критерием.

5. Почему план А-В-А-В имеет преимущество перед планом А-В-А?

1) он не имеет преимуществ; план А-В-А предпочтителен, так как проще и занимает меньше времени;
2) потому, что он содержит отмену воздействия;
3) так как в его случае сравниваются условное подкрепление и его отсутствие;
4) потому, что в его случае эффективность воздействия оценивается не один раз, а два.

Короткие эссе

1. Опишите проведенное Дресслером исследование лицевого зрения, расскажите, как были представлены результаты, и объясните, почему ученый использовал трех участников, а не одного.
2. Объясните, почему исследование Торндайка является хорошей иллюстрацией принципа простоты.
3. На примере эксперимента по научению дифференциации покажите, как обобщение данных может скрыть действия участников во время исследования.
4. На примере поведения крысы, нажимающей рычаг в проблемном ящике Скиннера, проиллюстрируйте утверждение этого ученого о возможности предсказания и управления поведением, если известны три основных фактора.
5 Что такое режим подкрепления? Приведите два примера.
6. Считается, что работы Скиннера отражают «технологический идеал». Объясните, почему.
7. Опишите три важнейших элемента каждого из планов с одним субъектом.
8. Опишите основные особенности плана с отменой и охарактеризуйте отличия планов А-В-А, А-В-А-В и А-В-С-В.
9. Дайте определение плана с несколькими базовыми уровнями, укажите, в каких случаях он имеет преимущество перед планом с отменой, и опишите три его разновидности.
10. На примере исследования мальчиков, страдающих и не страдающих ожирением, опишите основные особенности плана с изменяющимся критерием. В ответе используйте понятие «формирования».
11. Опишите три направления критики планов с одним субъектом.
12. Охарактеризуйте разницу между абсолютным и дифференциальным порогами и опишите три метода определения порогов, предложенных Фехнером.

Упражнения

Упражнение 11.1. Разработка программ саморазвития

Разработайте программу с изменяющимся критерием для одного из перечисленных ниже проектов саморазвития. Для каждого из проектов дайте операциональное определение целевого поведения или нескольких вариантов поведения, решите, что бы вы стали использовать в качестве подкрепления, и укажите критерий успеха.

1. Увеличение времени эффективного обучения.
2. Разработка программы физических упражнений.
3. Переход к более здоровому питанию.
4. Совершенствование организации времени.

Упражнение 11.2. Гипотетические результаты прикладного анализа поведения

Для каждого из следующих планов начертите график для одного испытуемого, иллюстрирующий противоположные результаты.

1. ПланA-B-C-B:
   1) подкрепление действует, но только если оно появляется при наличии определенного поведения;
   2) подкрепление действует вне зависимости от условий его появления.
2. План A-A,-B-A,-B:
   1) лекарство действует;
   2) неизвестно, действует ли лекарство; возможен эффект плацебо.
3. План с несколькими базовыми уровнями для трех различных условий:
   1) программа действует;
   2) нельзя не учитывать эффекты истории и взросления.
4. ПланA-B-A-B:
   1) программа действует;
   2) трудно сказать, вызваны ли изменения программой или другим фактором, например взрослением.

Упражнение 11.3. Графическое изображение результатов прикладного анализа поведения

Для каждого из приведенных ниже описаний и наборов данных начертите график для одного субъекта, точно отображающий результаты. Напишите краткие выводы.

1. Для снижения количества пауз на уроке во втором классе, вызванных поведением нарушителей дисциплины, был использован план A-B-A-B. Воздействие состояло в том, что учитель игнорировал моменты нарушения дисциплины со стороны ребенка и уделял ему особое внимание, если он действовал

продуктивно (например, выполнял задания). Количество пауз, зафиксированных за 1 час, было следующим:

1) во время первого А: 12, 12,7,6,6,9,8, 10,9, 11;
2) во время первого В: 9, 8, 9, 4, 3, 2, 2, 1, 4, 2;
3) во время второго А: 4, 5. 10, 6, 12, 10, 10, 10, 12, 9;
4) во время второго В: 9, 9, 2, 1, 1, 1, 0, 3, 4,1.

2. На тренировках у баскетболистов для повышения количества попаданий при свободных бросках использовался план с несколькими базовыми уровнями для нескольких человек. За успешные броски спортсмены получали очки, которые далее они могли обменять на более весомые награды. Приведенные ниже цифры обозначают количество удавшихся свободных бросков из 50 попыток. Подчеркнутые цифры представляют собой базовый уровень.

   Игрок 1:32.29.38.31.33. 44, 36, 37, 44, 41, 40, 38, 45, 42, 40, 44.

   Игрок 2: 30.32.28.30.30.40.35.32.33. 38, 40, 45, 44, 44, 42, 44.

   Игрок 3: 22. 28. 29. 28. 26. 25. 22. 26. 21.21. 23. 24. 35, 39, 40, 39.

ГЛАВА 12

Описательные методы исследований

Обзор задач главы

Последняя глава этой книги посвящена методам психологических исследований, которые в целом можно назвать описательными. Основная цель таких исследований — дать точное описание взаимоотношений, возникающих между индивидуумом и окружающей средой. В некоторых обстоятельствах описательный подход является единственно приемлемым, но в основном описательные исследования служат богатым источником гипотез для дальнейших изысканий. В них часто используется качественный анализ изучаемых явлений, но также может проводиться и сложный количественный анализ. В данной главе будут рассматриваться следующие методы: наблюдение, опрос, исследование конкретных примеров и архивное исследование. Прочитав эту главу, вы:

- поймете разницу между естественным наблюдением и наблюдением участников исследования;
- сможете назвать проблемы, связанные с наблюдением (например, проблемы контроля, этические проблемы, проблемы искажения и реактивности), и способы решения этих проблем;
- сумеете описать три вида опросов и указать достоинства и недостатки каждого из них;
- будете в состоянии объяснить проблемы (например, искажения, вызванные социальным ожиданием или формулировкой вопросов), затрудняющие интерпретацию данных, полученных в ходе опроса;
- сможете описать два вида исследований конкретных примеров и оценить адекватность каждого вида исследований;
- сможете описать важнейшие особенности архивных исследований и оценить этот исследовательский подход.

Как видно из структуры этой книги, центральное место в ней занимает экспериментальный метод. Ему были посвящены главы с 5 по 8, а предыдущие главы (в особенности глава 3) служили подготовкой к его изучению. Как вы помните, роль экспериментального метода определяется тем, что он дает возможность делать выводы о причинах и следствиях изучаемого явления, по крайней мере в том смысле,

как понимают причинность психологи. В последних трех главах вы познакомились с методами, которым некоторые блюстители традиций могут отвести лишь второстепенную роль, поскольку выводы, которые можно сделать в ходе исследований с применением этих методов, менее надежны, чем выводы па основе эксперимента. Но несмотря па то, что на основании корреляционных и квазпэксперпментальных исследований, а также исследований с малым *N* довольно сложно сделать выводы о причинах и следствиях, для полного понимания закономерностей поведения такие исследования просто необходимы. То же самое можно сказать и о методах, с которыми вы познакомитесь в этой главе. Все они имеют свои слабые стороны, но при этом являются неотъемлемой частью научного аппарата.

Наблюдение

Как вы узнали из главы 1, в которой рассматриваются цели психологии, поведение невозможно предсказывать и объяснять, им нельзя управлять без предварительного точного описания этого поведения. Главная цель исследований-наблюдений заключается в получении подобной описательной информации. Данный вид исследований позволяет подробно описать естественное поведение индивидуумов или групп индивидуумов в определенных условиях.

Виды наблюдений

Наблюдения, проводимые с целью описания поведения, можно классифицировать по двум параметрам. Во-первых, они различаются по степени вовлеченности экспериментатора в контакт с участниками исследования. Иногда исследователь устанавливает лишь очень поверхностные отношения с наблюдаемой группой, в других случаях он может даже стать ее членом. Таким образом, этот вид исследований включает соответственно «естественные наблюдения» и «наблюдения участников исследования». Во-вторых, наблюдения могут различаться в зависимости от степени воздействия экспериментатора на наблюдаемые условия. Этот показатель может варьировать от нулевого значения, когда исследователь принимает уже существующие условия и просто наблюдает, не вмешиваясь в происходящее, до достаточно высокого, когда он сам создает определенные условия, а затем наблюдает, что в них происходит. И естественные наблюдения и, наблюдения участников исследования могут проходить в условиях более или менее естественных.

Естественное наблюдение

При естественном наблюдении изучается поведение людей или животных в естественных для них условиях. Полигоном для естественных наблюдений могут быть и Детский сад, и африканские тропические леса, а объектами наблюдения могут быть люди всех возрастов и животные практически любых видов. В некоторых случаях не вполне естественная ситуация является достаточно пригодной для проведения естественного наблюдения. Пример тому — изучение поведения животных в современных зоопарках, довольно точно воспроизводящих нормальные условия их обитания.

В некоторых естественных исследованиях наблюдатель скрыт от наблюдаемых испытуемых. К примеру, при изучении способности дошкольников делиться чем-**либо** с другими детьми наблюдатель может находиться в соседней комнате и ви-

деть детей сквозь прозрачное зеркало (наблюдатель смотрит сквозь него, а дети видят только зеркало). В торговых центрах наблюдатель, изучающий ритуалы знакомств у городских подростков, может просто сидеть на скамейке в стратегически выгодном месте. В других исследованиях наблюдатель может вообще отсутствовать: все большее количество естественных наблюдений проводится с использованием записывающей аппаратуры. Потом сделанные записи просматриваются и исследователь оценивает интересующее его поведение.

При естественном наблюдении животных наблюдателю не всегда удается скрыть свое присутствие, так как животные хорошо чувствуют постороннего. При таких обстоятельствах наблюдатель обычно и не пытается прятаться и надеется, что через некоторое время животные привыкнут к нему и станут вести себя естественно. Такой подход позаимствован из антропологии: ученые-антропологи подолгу живут вместе с членами изучаемого удаленного от цивилизации племени.

Наблюдение участников исследования

Иногда исследователь сам входит в наблюдаемую группу — в таком случае исследование называется **наблюдением участников исследования.** Главное достоинство такого подхода заключается в том, что исследователь максимально приближен к происходящему. Вхождение в наблюдаемую группу дает исследователю непосредственный опыт, которого не имеет посторонний наблюдатель.

Психологической классикой стало одно из исследований, проведенное как наблюдение участников. Небольшая группа социальных психологов присоединилась к религиозной общине, чтобы изучить присущий ее членам образ мысли. В частности, они сформулировали следующий эмпирический вопрос: «Каковы ваши действия, если вы публично предсказываете конец света, а он не наступает?» Ответ был настолько удивительным, что на его основании была разработана одна из самых известных теорий социальной психологии — теория когнитивного диссонанса. Во вставке 12.1 вы найдете более подробную информацию об исследовании, посвященном судьбе неудачных предсказаний.

Оценка наблюдений

Исследователь, проводящий наблюдения, должен быть готов к возникновению некоторых проблем, в том числе проблемы недостаточного контроля, к появлению внесенного наблюдателем искажения, к проблеме реактивности, а также к этическим проблемам, связанным с вторжением в личную жизнь и процедурой получения осведомленного согласия.

Отсутствие контроля

Определенная степень контроля над ситуацией может быть достигнута, если наблюдатель имеет возможность на нее воздействовать (см. пример 32), но, как правило, исследователи должны довольствоваться тем, что предлагают обстоятельства. Поэтому основанные на наблюдениях выводы должны делаться с большой осторожностью. Если ребенок *A* подбирает игрушку, и сразу вслед за этим ребенок *B* подбирает точно такую же игрушку, означает ли это, что *B* подражает *A?* Может быть, это именно так, но возможны и другие объяснения, например эта игрушка привлекательна. При изучении описаний наблюдений вы не встретите высказываний типа «X вызывает 7».

ВСТАВКА 12.1

Классические исследования — когда предсказания не сбываются

Одной из ведущих теорий в социальной психологии в 60-70-х гг. XX в. была теория когнитивного диссонанса, разработанная Леоном Фестингером. Согласно этой теории, если мы пытаемся совместить противоречащие друг другу идеи, мы чувствуем дискомфорт (т. е. переживаем когнитивный диссонанс) и стараемся снизить этот диссонанс, убеждая себя в том, что все прекрасно и никакой проблемы нет. Одно из следствий данной теории состоит в том, что если мы прилагаем к чему-то большие усилия, а результат не соответствует нашим ожиданиям, то мы должны будем убедить себя в том, что усилия были приложены не зря. Фестингер получил возможность проверить свое предположение, встретив статью со следующим заголовком (Festinger, Riecken & Schachter, 1956, p. 30): «Планеты предсказывают. Кларион взывает к городу: спасайтесь от наводнения. Нас затопит 21 декабря, сообщил космос жительнице пригорода».

В статье рассказывалось о том, что миссис Китч предсказала наводнение, которое в конце декабря разрушит большую часть Северной Америки. Откуда она это узнала? Миссис Китч утверждала, что вступила в непосредственный контакт с жителями планеты Кларион, прилетавшими на Землю на летающих тарелках. Они якобы видели линии разломов коры, из которых должны исторгнуться огромные водные массы. Миссис Китч собрала небольшую группу последователей, что-то вроде религиозной общины, которые должны были убедить мир покаяться прежде, чем станет слишком поздно (а время уходило; было предсказано, что катастрофа произойдет через 4 месяца).

Фестингер предположил, что наводнения не будет, и заинтересовался тем, как поведут себя миссис Китч и ее группа, когда предсказание не сбудется. Он решил увидеть все своими глазами. Пригласив двух коллег и наняв пять наблюдателей, ученый на несколько недель присоединился к религиозной группе, став одновременно участником исследования и наблюдателем. То, что они обнаружили, описано в замечательной книге (Festinger et al., 1956). Она называется так же, как эта вставка. С методологической точки зрения важны несколько моментов.

Во-первых, возникли трудности с записью информации. Наблюдатели не хотели обнаруживать свои истинные цели и не вели записи во время собраний в доме миссис Китч. Поэтому им пришлось полагаться на память больше, чем хотелось бы. Затруднение было в некоторой степени преодолено, когда исследователи догадались использовать для ведения записей ванную комнату. Второй проблемой была реактивность. Их истинные цели вряд ли были известны группе. Но исследователи беспокоились, что их присутствие подкрепляет убеждения ее членов. За короткий период в группу вошло так много новых людей, что миссис Китч сочла (а следовательно, так решила и ее группа) это «знаком свыше» того, что предсказание о наводнении верно. Группа думала, что двое из наблюдателей «посланы Хранителями» (p. 242) — жителями Клариона, а один наблюдатель *сам прилетел с этой планеты.* Возможно, что наблюдатели поэтому в какой-то мере укрепляли веру членов группы миссис Китч, но действие этого фактора было трудно оценить. В данном случае фактор реактивности создает и этическую проблему: то, что присутствие Фестингера и его коллег укрепляло убеждения членов группы, может быть расценено как их вклад в развитие «патологического состояния».

Поскольку вы читаете эту книгу, вы знаете что 21 декабря 1950 г. конца света не произошло. Это было не единственное несбывшееся предсказание. Миссис Китч сказала группе, что Хранители пришлют космический корабль, который 17 декабря приземлится у нее во дворе

и спасет их. Корабль, конечно, не прибыл, а 4 дня спустя не наступил конец света. Возможно, группа была обескуражена, бросила свою деятельность, назвала миссис Китч сумасшедшей и вернулась к нормальной жизни? Нет. Большинство членов группы стали еще *более активно* пытаться обращать людей в свою веру. Очевидно, чтобы оправдать свои усилия, они убедили себя в том, что их деятельность предотвратила катастрофу - группа «распространила так много света, что Бог спас мир от разрушения» (р. 169). Так, несбывшееся предсказание может не изменить, а укрепить взгляды человека.

Наблюдения могут также служить целям фальсификации — очень важной стратегии, используемой для проверки теории. Наблюдение, согласующееся с теоретическими предположениями, поддерживает их по индукции, но наблюдения, противоречащие теории, могут оказаться даже более информативными. Как вы помните из главы 3, одно несоответствие еще не опровергает теорию, но, без сомнения, ставит ее под вопрос. Например, влиятельная в 60-х гг. теория агрессивности животных (Lornz, 1966) утверждала, что битвы между представителями одного вида практически никогда не приводят к смерти участников (за исключением битв между людьми). Годами специалисты спорили о том, почему агрессивность людей так сильно, и причем в худшую сторону, отличается от агрессивности животных. Сама постановка вопроса о выяснении причин различий предполагала, что эти различия действительно существуют, но проведенные наблюдения поставили под сомнение факт отсутствия летальных исходов в стычках животных. Например, при борьбе шимпанзе за территорию исследователи наблюдали поведение, похожее на инциденты на границах (Goodall, 1978). Если несколько обезьян из одной группы встречают бродящее в одиночку животное пз другой группы, то они практически наверняка нападут и убьют его. Это открытие ставит под вопрос утверждение Лоренца о том, что проявления агрессии у животных редко приводят к летальному исходу.

Более поздний пример — исследование Боша-Акерманна и Боша (Boesch-Achermann & Boesh, 1993). Они наблюдали, как шимпанзе обучают своих детенышей использовать орудия труда (действия с молотком и наковальней), чтобы раскалывать различные виды орехов (см. рис. 12.1). Это важное исследование вызвало вопросы по поводу сложившихся ранее представлений. Авторы представили это следующим образом:

В последних критических обзорах исследований процесса обучения у животных отрицалось, что животные обладают способностью к подражанию, но наблюдаемые нами примеры процедуры обучения были бы лишены всякой функциональности, если бы шимпанзе не обладали способностью к подражанию. Многие считают, что педагогика — это исключительно человеческое явление, по наши наблюдения шимпанзе этого не подтвердили.

Несмотря на трудности, связанные с контролем, наблюдения могут дать важную и полезную информацию. Они могут поставить существующие идеи под сомнение, а также стать важным источником гипотез для дальнейших исследований. Так, исследование Боша-Акерманна и Боша вызвало сомнение в правильности существующих представлений о способности шимпанзе к обучению себе подобных, а также создало возможность для дальнейших исследований этой способности приматов. Например, могут ли шимпанзе, живущие в неволе и обученные операциям с орудиями труда, научить детенышей выполнять эти операции?



Рис. 12.1. Шимпанзе, использующие орудия труда. Взрослая самка ест орех, который она только что расколола каменным молотком, а ее сын изучает скорлупу. Самки шимпанзе делятся орехами со своими детенышами, тем самым обучая их использовать орудия для раскалывания орехов (из книги Boesch-Achermann & Boesh, 1993)

Искажение, вызванное наблюдателем

Вторая проблема наблюдений — это искажения, вносимые наблюдателем. Из главы 6 вы узнали, что если экспериментатор ожидает определенных результатов, то само его ожидание может повлиять на ход исследования. При наблюдении **искажение, вызванное наблюдателем,** означает наличие предвзятого мнения об объекте, наблюдение через призму этого мнения. Рассмотрим, к примеру, что может произойти, если исследователь изучает агрессию у дошкольников и заранее уверен в том, что мальчики более агрессивны, чем девочки. Такой наблюдатель оценит одно и то же неоднозначное поведение как агрессивное у мальчиков и как нормальное у девочек. Аналогично при исследовании животных наблюдатели, чьи мнения о способности животных к альтруизму расходятся, могут по-разному интерпретировать неоднозначное поведение. Кроме того, искажение может появиться из-за огромного количества собранной информации. Концентрация наблюдения на конкретной задаче снижает объем сбора информации, но на отбор данных при этом могут подействовать уже имеющиеся убеждения.

Искажение может быть заметно снижено, если использовать четкие операциональные определения и предварительно точно сориентировать наблюдателей на определенный вид поведения. Для проведения наблюдений обычно используются **контрольные таблицы поведения.** Такая таблица представляет собой набор определений различных видов поведения, которые должны уметь различать наблюдатели. Рассмотрим, к примеру, исследование, в ходе которого наблюдалось поведение доставщиков пиццы. У этой группы количество аварий на дорогах в три раза превышало средний показатель по стране. Людвиг и Геллер (Ludwig & Geller, 1997) использовали контрольную таблицу «разработанную на основе десятилетних наблюдений за поведением водителей и двухлетних — за доставщиками пиццы» (р. 254). Таблица включала наблюдения за поведением водителей при повороте на

главную дорогу на перекрестке рядом с пиццерией. На перекрестке находился знак «Стоп». Поведение водителей на этом месте разделили на три категории: полная остановка, медленное движение вперед и быстрое движение вперед. Каждая из категорий имела свое определение. Например, медленное движение вперед было определено как движение автомобиля мимо знака «стоп» со «скоростью, примерно равной скорости движения взрослого человека» (р. 255). Наблюдатели также отмечали ситуацию на дороге, включал ли водитель указатель поворота и был ли пристегнут ремень.

Для контроля искажения, вызванного наблюдателем, можно также использовать сравнение данных нескольких наблюдателей. Этот способ гарантирует одни из видов надежности опыта — межнаблюдательскую надежность (согласованность наблюдения), что обычно определяют как процент случаев согласия наблюдателей друг с другом. Конечно, разные наблюдатели могут вносить искажения в одинаковой степени, но одновременное использование контрольных таблиц, обучения наблюдателей и согласования их данных обычно позволяет снять искажение. Кроме того, искажение можно снизить автоматизацией процедуры наблюдения. Мы уже встречались с ситуацией, когда видеозапись позволяла увеличить объективность исследования.

Кроме того, искажение можно снизить, если использовать различные процедуры выборки информации. Например, при наблюдении иногда используют **временную выборку.** В соответствии с этой процедурой запись событий ведется только в заранее определенные промежутки времени. Такие промежутки можно отобрать по определенному правилу или случайным образом. Событийная выборка направлена на отбор для наблюдения определенных событий, все остальные события игнорируются. Эта процедура использована в примере 31.

Реактивность испытуемых

Подумайте о всех тех вещах, которые вы обычно делаете в течение дня. Делали бы вы их точно так же, если бы знали, что за вами наблюдают? Вероятно, нет. По всей видимости, вы проявили бы реактивность. Это значит, что ваше поведение изменилось бы, если бы вы знали, что его фиксируют. Очевидно, что такая проблема легко может возникнуть при проведении исследований — вот почему получили широкое распространение такие устройства, как прозрачное зеркало. Эта проблема также возникает при исследованиях животных, если наблюдатель не имеет возможности спрятаться. Исследователи полагают, что через некоторое время животные привыкают к присутствию постороннего, но в какой мере — трудно оценить.

Реактивность можно снизить проведением незаметных измерений. Это любые измерения поведения, прямые или косвенные, при условии, что субъекты не знают об их проведении. Прямые незаметные измерения включают скрытую видео- или аудиозапись, а также фиксирование событий и их результатов скрытым наблюдателем. Косвенные незаметные измерения направлены на события и их последствия, предположительно вызванные определенным поведением, которое само по себе не наблюдается. Вебб с соавторами (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981) описали ряд объектов для косвенных измерений — некоторые из них весьма оригинальны. Приведем примеры нескольких таких объектов:

- содержимое мусорного ведра при изучении особенностей питания;
- пыль на книгах как показатель частоты их использования;
- степень износа коврового покрытия, расположенного в стратегически важном месте, при изучения особенностей передвижения людей;
- анализ политических наклеек для бамперов в год выборов.

Этика

Я думаю, вы уже понимаете, что снижение реактивности неизбежно приводит к возникновению этической проблемы вторжения в личную жизнь и осложняет процедуру обоснованного согласия. Понравится ли вам, если вы обнаружите, что исследователь прячется **у** вас в спальне под кроватью и записывает все, что вы говорите и делаете? Верьте или нет, по такое исследование было проведено (Henle & Hubbell, 1938), хотя в настоящее время оно не получило бы разрешения Исследовательского совета.

Во вставке 3.1 я указал, что некоторые исследователи не хотят проводить полевые исследования, поскольку беспокоятся о возможном нарушении права на личную жизнь. Но этический кодекс *АРА (АРА,* 1992) разрешает проведение наблюдений при условии, что будут приняты определенные меры предосторожности. Например, получение обоснованного согласия участников не считается необходимым, если поведение изучается в общественном месте (а не в чьей-то спальне), люди не подвергаются никакому воздействию и сохраняется конфиденциальность информации. Связанные с наблюдениями стандарты приведены в табл. 12.1.

Таблица 12.1

Этический кодекс *АРА* и наблюдения

Приведенные ниже стандарты этического кодекса *АРА* 1992 г. особенно важны для наблюдений и являются дополнением к обычным нормативам исследований.

Стандарт 1.14. Непричинение вреда

Психологи делают все возможное, чтобы избежать причинения вреда... участникам исследования... и снизить предполагаемый ущерб, если его нельзя избежать.

Стандарт 5.03. Сведение к минимуму вторжения в личную жизнь

а) Чтобы снести к минимуму вторжение в личную жизнь, психологи включают в письменные и устные отчеты, консультации и др. только информацию, соответствующую целям составления этих отчетов, проведения консультаций и др.

б) Психологи обсуждают конфиденциальную информацию... или оценивают данные, касающиеся участников исследования... преследуя исключительно научные или профессиональные цели и только со специалистами по данному вопросу.

Стандарт 5.08. **Использование конфиденциальной** информации **с** дидактическими или другими **целями**

а) Психологи в своих работах, лекциях или другим адресованным обществу путем не раскрывают конфиденциальные данные об...участниках своих исследовании... если последние не дали на это письменного согласия или не получено другое этическое или юридическое разрешение.

б) В обычных случаях при подобной научной и профессиональной деятельности психологи скрывают персональную информацию, касающуюся участников исследования... чтобы устранить возможность их идентификации^

Стандарт 6.12. Освобождение от обоснованного согласия

Прежде чем решить, что планируемое исследование (включающее, например, анонимные опросы, естественные наблюдения или некоторые виды архивных исследований) истребует получения от участников обоснованного согласия, психологи изучают соответствующие инструкции, требования исследовательского совета и консультируются с коллегами.

Стандарт 6.13. Осведомленное согласие при аудио- или видеозаписи исследования

Психологи получают осведомленное согласие от участников исследования, прежде чем производить какую-либо видео-или аудиозапись участников, если только исследование не проводится в виде естестве!того наблюдения в общественном месте и не предполагается, что запись будет использована для идентификации участников или причинения им вреда.

Источник: American Psychological Association, 1992.

Обратите внимание, что в двух пунктах таблицы, в стандартах 6.12 и 6.13, говорится о том, что при определенных обстоятельствах проведение естественных наблюдений не требует получения осведомленного согласия. Но что если наблюдение проводит участник исследования? У вас может вызвать вопрос отсутствие обоснованного согласия в исследовании Фестиигера (вставка 12.1). Ученый имел все основания считать, что если будет использована обычная процедура, он и его коллеги не смогут войти в группу. Вопрос в том, как сегодня отнесся бы исследовательский совет к такому исследованию. Наблюдение участников исследования часто используют при проведении качественных исследований, и получение осведомленного согласия у наблюдаемой изнутри группы — это сегодня обычная процедура. Отказ от получения осведомленного согласия в наши дни требует строгого обоснования (Taylor & Bogdan, 1998).

В завершение раздела о методе наблюдения рассмотрим два исследования, которые иллюстрируют вторую из описанных выше категорий — меру естественности условий наблюдения. Первое исследование представляет собой чистое наблюдение без какого-либо прямого вмешательства со стороны исследователя; второе проведено в менее естественных условиях.

Пример 31. Естественное наблюдение прикосновений

Когда мужчины и женщины вместе находятся в общественном месте, то нередко можно наблюдать, как они прикасаются друг к другу. Кто чаще инициирует прикосновения, мужчины или женщины? Проявляются ли половые различия в типах прикосновений? Существуют ли в этом возрастные различия? Эти и другие вопросы в ходе естественного наблюдения поведения пар в общественных местах исследовали Халл и Веккиа(Ha11 & Veccia, 1990). Работа этих ученых хорошо демонстрирует трудности проведения естественных наблюдений.

Халл и Веккиа обучили пять наблюдателей подробно, но при этом незаметно фиксировать прикосновения лиц разного и одного пола в различных общественных местах: от аэропортов до метро и торговых центров в районе Бостона. Как и все исследователи, проводящие наблюдения, Халл и Веккиа прежде всего встретились с проблемой операциональных определений. Как точно определить прикосновение? В какой момент должна начинаться и заканчиваться запись такого события? Что конкретно должно фиксироваться при каждом наблюдении?

Для записи описании происходящего наблюдатели использовали портативные записывающие устройства и небольшие хронометры, издающие сигнал каждые 10 секунд. Чтобы окружающие не слышали сигналов хронометров (т. е. чтобы проводить наблюдение незаметно), наблюдатели пользовались наушниками. Интервал между сигналами использовался для записи информации о поведении конкретной пары. Во-первых, оценивался их возраст (с шагом в 10 лет); далее записывалось, прикасались ли они друг к другу в начале интервала, по ходу записи или вообще не прикасались. Если прикосновение было, записывалось, какие части тела соприкасались и каким образом (например, люди могли держаться за руки). В заключение, фиксировалось, кто из двух людей прикоснулся первым. Для разнополых пар прикосновения кодировались как инициированные мужчиной (МЖ) или женщиной (ЖМ). Это было не всегда возможно, особенно если прикосновение уже осуществилось к началу интервала, но межнаблюдательская надежность данного исследования была довольно высокой (91% согласий между наблюдателями).

Если участники передвигались (например, шли по магазину), наблюдатели отбирали пары с помощью временной выборки. Выбрав пару, они фиксировали ее поведение в течение 10 секунд, 10 секунд ждали, а затем выбирали новую пару. Также использовалась определенная разновидность событийной выборки: по истечении 10 секунд ожидания отбирали пару, ближайшую к заранее выбранному месту.

На основании проведенных ранее исследований Халл и Веккиа полагали, что в разнополых парах инициировать прикосновения будут в основном мужчины, поскольку у них желание быть сильным и контролировать ситуацию выше, чем у женщин. К своему удивлению ученые не обнаружили никаких различий: в целом, количество прикосновений МЖ равнялось количеству прикосновений ЖМ. Но более тщательный анализ дал интересные результаты. Например, данные из табл. 12.2 указывают на сильное влияния возраста. У более молодых пар, особенно у людей младше 30 лет, чаще наблюдались прикосновения МЖ. У старших пар наблюдалось обратное: женщины чаще инициировали прикосновение. Данное явление можно объяснять по-разному, но Холл и Веккиа предположили, что мужчины чувствуют более сильную необходимость в утверждении своей силы и способности контролировать ситуацию на ранних этапах отношений, а отношения взрослых пар обычно более давние. (Конечно, возможно, что у более взрослых пар прикосновения ЖМ типа «локтем в бок» вызваны задумчивыми взглядами, которыми мужчина провожает обнимающиеся юные пары.)

Таблица 12.2

Возрастные различия в прикосновениях у разнополых пар

	Кол-во прикосновений		Процент прикосновений	
Возраст пары	МЖ	ЖМ	МЖ	ЖМ
Подростки	16	6	73	27
20-30 лет	58	46	56	44
30-40 лет	28	29	49	51
40+	7	26	21	79

Примечание. МЖ - прикосновение, инициированное мужчиной; ЖМ = прикосновение, инициированное женщиной.
Источник: Адаптировано из Hall & Veccia, 1990, tabl. 6.

Также Халл и Веккиа обнаружили интересное различие типов прикосновений. Хотя при прикосновениях рук не было выявлено никаких значимых половых различий, были обнаружены различия в прикосновениях, обозначенных как «обнимание рукой» и «сцепление рук». Мужчины чаще демонстрировали первое, а женщины — последнее.

Пример 32. Наблюдение альтруизма у детей

Исследование Петерсона с соавторами (Peterson, Ridley-Johnson & Carter, 1984) представляет собой интересный пример естественного наблюдения с элементами искусственно созданной ситуации. Ученые исследовали готовность школьников к оказанию помощи. Предшествующие наблюдения показали, что такое поведение возникает спонтанно и не чаще чем один-два раза в час. Низкая частота его появления навела исследователей на мысль о создании особой ситуации, в которой подобное поведение должно проявиться с большой степенью вероятности. Таким образом, одна из причин введения элементов искусственности в ситуацию наблюдения состояла в увеличении частоты этих, вообще довольно редких событий.

Детям сказали, что каждый из них по очереди будет носить «суперплащ», «ярко-синюю накидку с красной атласной звездой на груди» (Peterson et al. 1984, р. 237). Накидка застегивалась сзади па шее большой пуговицей. Предварительная проверка показала, что для ее застегивания необходима помощь другого человека. На этапе свободной игры дети по очереди носили накидку; очередь при этом определялась вытягиванием бумажек с именами из банки. Каждый ребенок носил суперплащ около 4 минут.

Происходящее записывалось с помощью видеокамеры, а затем два независимых наблюдателя просматривали записьи оценивали помощь детей друг другу. Межнаблюдательская надежность была высокой: процент согласия был более 90%. По схеме кодирования, отображающей операциональные определения типов поведения, были выбраны две основные категории: «донорское» поведение, демонстрируемое потенциальными помощниками, и «реципиентное» поведение носителя плаща. Примером донорского поведения была *«спонтанная помощь* — когда без словесных или жестовых указаний донор словами или действиями предлагает помочь застегнуть пуговицу на плаще» (р. 237; курсив в оригинале). Примером рсципиентного поведения была *«просьба о помощи* — когда реципиент словесно просит о помощи... или подходит к донору, поворачиваясь так, чтобы пуговица была видна потенциальному донору, и жестами просит помочь» (р. 238; курсив в оригинале).

Каковы были результаты? Из 56 возможных случаев дети оказали спонтанную помощь 32 раза, и 13 раз — помощь после напоминания. Но самым интересным оказалась реакция реципиентов на оказание помощи. К своему удивлению, исследователи обнаружили, что дети редко благодарили за помощь. По сути «большинство детей-реципиентов на помощь реагировали скорее негативно (например, говорили "убирайся" и толкались в ответ на попытку другого ребенка помочь застегнуть плащ), чем позитивно» (р. 238)! Редко наблюдалась обоюдная помощь (ребенок *А* помогает ребенку *В,* а позднее ребенок *В* помогает ребенку *А).*

Кроме того, исследователи попросили каждого ребенка оценить социальную компетентность одноклассников и выяснили, что в основном помощь оказывали

дети, у которых этот показатель был высоким. Данная процедура оценки показывает, что наблюдение может включать элементы других методов, в данном случае — определенный вид опроса. Таким образом, мы подошли к рассмотрению следующей темы.

Опрос

Метод основывается на простой идее: если вы хотите узнать, что думают люди на определенную тему, следует просто спросить их об этом. **Опрос** — это структурированный сборник вопросов пли утверждений, который выдается группе людей для выявления их отношения к чему-либо, их убеждении, ценностей пли особенностей поведения. За годы применения метода была проведена оценка самых разнообразных характеристик, от политических взглядов до особенностей сексуальной жизни.

В противоположность большинству методов, о которых мы рассказывали в этой книге, опрос требует уделять особое внимание процедуре выборки. Как вы помните из главы 4, иногда исследователи пытаются оценить определенные особенности (например, политические предпочтения) большой категории людей — популяции — на основании данных, собранных у небольшой группы — выборки. Чтобы результаты опроса были валидными, выборка должна отражать особенности популяции, т. е. должна быть **репрезентативной.** Таким образом, в методе опроса адекватность сделанных выводов зависит от качества процедуры отбора участников.

Виды опросов

Есть три основных способа сбора данных в ходе опроса. Каждый из них имеет свои сильные и слабые стороны, касающиеся выборок и других аспектов. При проведении опросов для сбора данных используют непосредственные, персональные интервью, письменные анкеты, телефонные интервью или одновременно несколько методов.

Интервью

Вы, вероятно, слышали о докладе Кинзи — пожалуй, самом известном американском опросе на сексуальную тему. Опрос, состоящий из подробных персональных интервью тысяч мужчин и женщин, был проведен после Второй мировой войны. По его результатам были составлены две большие книги о сексуальном поведении людей в Америке — одна о мужчинах (Kinsey, Pomeroy & Martin, 1948), а вторая о женщинах (Kinsey, Pomeroy, Martin & Gebhard, 1953). Вы можете предположить, что использованный Кинзи метод интервью, не позволил узнать у опрашиваемых интимные подробности их сексуального поведения π существующих в этой сфере установок, особенно если принять во внимание историческую эпоху, в которую проводилось данное исследование. В действительности консервативная послевоенная Америка была шокирована многочисленными откровениями о добрачной половой жизни, мастурбации и прелюбодеяниях. Хотя книги были написаны в сухой академической манере и были заполнены таблицами и гистограммами, они стали бестселлерами и вызвали многочисленные споры по поводу личности Кинзи.

Одни обвиняли его в том, что он способствует нравственному упадку общества и даже является коммунистом, а другие считали его пионером научных исследований одного из важных аспектов поведения людей (Christenson, 1971).

Преимущество метода интервью заключается в его многосторонности и возможности получить весьма подробную информацию. Хотя обычно опрашивающий задает стандартный набор вопросов, опытный интервьюер способен извлечь довольно много информации, задавая дополнительные вопросы и делая уточнения. Присутствие интервьюера снимает проблему неясных вопросов, так как он всегда может объяснить непонятное. (Иногда возникают проблемы с отбором участников, поскольку возможность проинтервьюировать значительные части популяции зачастую отсутствует, так как люди отказываются принимать участие в исследовании, их невозможно найти или они живут в таких районах, которых исследователи стараются избегать. Например, нищие очень часто плохо представлены в национальных опросах, проводящихся методом интервью.) Часто также проводятся групповые интервью, например интервью фокус-групп, о которых рассказывалось в главе 10.

С интервью связаны и другие проблемы — затраты на проведение, транспортировка и искажение, вызываемое интервьюером. Необходимо нанять и обучить интервьюеров, дорожные расходы могут быть весьма ощутимыми, а кроме того, такой опрос может быть сильно ограничен территориально из-за трудностей перемещения интервьюеров на большие расстояния. И несмотря на проводимое обучение, всегда есть возможность, что на даваемые с глазу на глаз ответы повлияет искажение, вызванное человеком, проводящим интервью. К примеру, может возникнуть межрасовое искажение, приводящее к систематическим расхождениям в ответах представителей расы интервьюера и других рас.

Внимательные исследователи разрабатывают программу подготовки, чтобы по возможности унифицировать процесс интервью. Кроме того, иногда для различных видов задач отбираются разные интервьюеры. Например, при опросе пенсионерок женщина среднего возраста может получить больше информации, чем молодая девушка, которую могут даже не пустить в дом (van Kammen & Stouthamer-Loeber, 1998).

Письменный опрос

Второй вид опроса — **письменный опрос.** Существует несколько его разновидностей, но обычно он состоит из перечня утверждений (участники отвечают, согласны они или нет) пли вопросов, на которые дается ответ. Вопросы могут быть либо открытыми, либо закрытыми. **Открытый вопрос** предполагает развернутый ответ, включающий некоторую описательную информацию. На **закрытый вопрос** можно ответить «да» или «нет» или выбрать один из представленных вариантов ответа. Чтобы понять разницу между этими типами вопросов, рассмотрим два варианта вопроса о школьном финансировании:

Открытый: Как, по Вашему мнению, необходимо финансировать общее образование?

Закрытглй: Считаете ли Вы, что налог па имущество должен идти на финансирование общего образования?

Для анализа результатов письменных опросов, содержащих закрытые вопросы, часто используют интервальную шкалу измерений. Как вы помните из главы 4, интервальная шкала не имеет точки истинного нуля, а интервалы между делениями шкалы одинаковы. Haute всего в опросах используется один из вариантов интервальной шкалы — шкала Лпкерта (носящая имя своего изобретателя, Ренсиса Ликерта). Стандартная шкала Лпкерта содержит от 5 до 9 точек, каждая из которых соответствует оценке из некоторого множества возможных. Например, при опросе отношения к телевизионному освещению новостей участников могут попросить оценить по 5-балльноп шкале, насколько они согласны с определенным набором утверждений. Утверждения могут быть следующими.

Человек, ведущий программу новостей, должен иметь не менее 5 лет опыта работы профессиональным журналистом.

КНС НС НУ С АС

Опрашиваемый обводит точку на шкале, соответствующую его убеждениям. КНС означает «категорически не согласен», НС — «не согласен», НУ — «не уверен» и т. д. Далее, точки можно преобразовать в 5-балльную шкалу (КНС = 1, НС = 2 и т. д.) и объединить ответы на разные вопросы по одной теме.

Как вы знаете, письменные опросы можно посылать по почте (возможно, вы уже занимались этим) или проводить в группах. Очевидно, что в последнем случае процент людей, заполнивших и вернувших опрос, будет выше. При использовании почты возврат опросников всегда является проблемой. Возврат, равный 85% или выше, — это великолепный результат (и очень редкий). 70-85% — это очень хорошо, а 60-70% — более или менее приемлемо (Mangione, 1998). Если возврат ниже 60%, то исследователи начинают беспокоиться о репрезентативности данных. Другая проблема с количеством возвратов возникает, если вернувшие заполненный опросник имеют какие-либо важные отличия от тех, кто не вернул. Такая проблема называется **искажением, вызванным отсутствием ответа** (Rogelberg & Luong, 1998). В случае такого искажения выводы о популяции в целом делать по меньшей мере рискованно.

При проведении опросов, рассылаемых по почте, неплохой процент возврата получают, если: а) опросник небольшой, его несложно заполнить и он состоит преимущественно из закрытых вопросов, а на открытые вопросы отвечать не обязательно; б) он начинается с относительно интересных вопросов, а более скучные (например, демографические) представлены в конце; в) прежде чем посылать вопросы, участников предупреждают о том, что к ним придет опросник и им будут весьма признательны за помощь; г) отсутствие ответа приводит к дополнительному напоминанию, повторному отправлению опросника и даже телефонному звонку с просьбой заполнить анкету; д) письмо с анкетой выглядит очень профессионально, подпись на письме сделана от руки, а ис распечатана на принтере или ксероксе и е) у опрашиваемого не возникает подозрений в том, что опрос — это шаг номер один в попытке продать ему что-либо. Также процент возврата можно увеличить, использовав небольшие денежные призы нли подарки (Fowler, 1993; Rogelberg & Luong, 1998). При достаточно высоком проценте возврата и правильно разработанных анкетах метод письменного опроса может принести ценную информацию и при этом быть более эффективным и требовать меньших затрат, чем интервью.

Телефонный опрос

Третий способ провести опрос — отобрать номера телефонов и обзвонить людей. Возможно, вы помните из главы 4, что выбор телефонов из телефонного справочника создал определенные проблемы для журнала «Литрери Дайджест» и он не смог сделать правильный прогноз о президентских выборах в 1936 г., поскольку, в то время не у всех были телефоны. Сегодня домашние телефоны имеют более 90% американцев. Но как же быть с номерами, не представленными в справочнике? Здесь может возникнуть проблема, так как но некоторым оценкам в списках отсутствует около трети номеров домашних телефонов (Lavrakas, 1998). Решить эту проблему можно с помощью процедуры, названной «генератор случайных чисел», в ходе которой простая компьютерная программа генерирует списки номеров телефонов, соответствующих определенной территории. Таким образом, каждый телефонный номер может попасть в выборку.

Очевидным преимуществом **телефонного опроса** является более низкий объем затрат на его проведение по сравнению с интервью и письменным опросом. Этот метод объединяет эффективность опроса, рассылаемого по почте, с возможностью личного контакта, свойственного интервью. По сравнению с интервью у данного метода имеются и другие преимущества — он позволяет опросить гораздо больше людей в единицу времени. Кроме того, телефонный опрос предотвращает возможность нападений на интервьюера, как это бывает при проведении полевых исследований в районах с высокой преступностью. Один из недостатков телефонного интервью состоит в том, что опрос должен быть кратким, так как люди могут потерять терпение и отказаться отвечать дальше; кроме того, это требует исключения из таких интервью открытых вопросов. Еще одна проблема связана с процентом возврата. Вероятно, вы думаете, что процент возврата у телефонных опросов довольно высокий. Но это не так. Во-первых, есть большое количество семей, в которых оба взрослых работают, а поэтому днем опрашивающий натолкнется чаще на автоответчик, чем на человека. Таким образом, многие звонки останутся без ответа, в том числе повторные звонки на один и тот же номер. Это привело к тому, что исследователи проводят телефонные опросы по вечерам. К сожалению, именно в это время звонят торговые агенты и предлагают потрясающе выгодно приобрести виниловое покрытие для обшивки дома или «бесплатный» отпуск в Орландо. Осложняет эту проблему то, что не знакомые с этикой продавцы в начале разговора пытаются выдать свой звонок за опрос. Один из способов предотвратить возможные недоразумения — это предварительно послать короткое письмо или открытку и предупредить о будущем звонке («пожалуйста, откликнетесь па наш звонок»).

Три варианта интервью, а также письменные и телефонные опросы используются не только по отдельности. В ходе интервью участникам могут устно задавать открытые вопросы, а затем попросить заполнить анкету, перед проведением телефонного опроса могут послать письменное уведомление и предварительное описание этого опроса, а за телефонным опросом могут последовать дополнительные интервью или письменные опросы. Каждый конкретный исследовательский проект предъявляет к исследователю свои требования. Иногда для достижения поставленных целей исследователь должен обладать определенным уровнем творческой активности. Хорошим примером комбинированной стратегии проведения иссле-

дования является групповое телефонное интервью, которое провела фирма, получившая заказ у издателя данной книги. В двух группах из восьми и десяти преподавателей, пользующихся вторым изданием этого учебника, провели две независимые телефонные конференции и выяснили мнение участников о том, как можно улучшить книгу перед третьим изданием. Хотя сначала я думал, что группы слишком велики, ведущий прекрасно выполнил свою работу, удерживал внимание людей на данной теме и опросил участников о сильных и слабых сторонах книги, а также о возможностях ее улучшения. Для меня обе записи интервью оказались очень полезными. Например, в ходе интервью мы решили создать руководство для « лабораторных занятий, и сегодня оно выходит вместе с учебником.

Оценка опросов

Опрос — это общепризнанный способ сбора данных, высоко ценимый за возможность получения важной информации о поведении людей. Но тем не менее при проведении опросов могут возникнуть серьезные проблемы. Первая из них касается построения выборки — эту проблему мы рассматривали выше. Очевидно, что неполноценная выборка может привести к получению неверных результатов. Вторая проблема заключается в возможном появлении искажения ответа, самым распространенным видом которого является **искажение, вызванное социальным ожиданием.** Иногда, в ходе опроса люди отвечают на вопросы не то, что они действительно думают или чувствуют, а то, что по их мнению, они должны отвечать. Другими словами, они пытаются создать положительный образ самих себя, образ, соответствующий социальным ожиданиям. В литературе по социальной психологии описано множество исследований, показавших, что мнение людей, высказываемое ими по некоторым вопросам, не всегда соответствует поведению этих людей. Поэтому результаты опроса должны интерпретироваться с учетом этого искажения, а выводы — подкрепляться результатами других исследований.

Третья важная проблема опросов касается их содержания. Вопросы и утверждения бывают весьма неоднозначными, как, например, в случае следующего утверждения:

Посещать родственников может быть весело.

Иногда составители опросов стараются включить в один вопрос как можно больше информации — это приводит к тому, что один вопрос может требовать двух ответов. Например, участников могут спросить, согласны ли они с таким утверждением:

Женщины в барах не должны сквернословить и покупать выпивку незнакомым мужчинам.

Участники опроса, которые согласны с утверждением о сквернословии, но не согласны с утверждением о покупке выпивки, не будут знать, что ответить на такой вопрос.

Кроме того, вопросы могут быть «наводящими» (этим термином активно пользуются адвокаты). Ниже представлены два варианта вопроса о Законе о чистом воздухе, на которые наверняка будут даны разные ответы:

Учитывая важность охраны окружающей среды для будущих поколений, как, по Вашему мнению, следует поступить с Законом о чистом воздухе, ужесточить, ослабить или оставить как есть?

Учитывая то, что установка газоочистителей может увеличить счета за электричество на 25%, как, по Вашему мнению, следует поступить с Законом о чистом воздухе, ужесточить, ослабить или оставить как есть?

Этот вид искажения часто возникает в мире бизнеса, когда опрос проводится с целью продажи какого-либо товара. Например, «Бургер Кинг», пытаясь превзойти «Макдоналдс» по продажам, провел опрос и заявил, что около 75% опрошенных предпочитают их «вупресы» «биг-макам». Ключевой вопрос они сформулировали следующим образом (Kimmel, 1996, p. 195):

Какие гамбургеры Вам больше нравятся, подрумяненные на открытом огие или жареные?

Понимаете, в чем здесь проблема? Конечно, бургеркинговское «подрумянивание на открытом огне» звучит более аппетитно, чем макдоналдсовская «жарка», поэтому результат опроса совсем неудивителен. Но в ходе другого опроса вопрос был задан по-другому (Kimmel, 1996, p. 195):

Какие гамбургеры Вам больше правятся, поджаренные па горячем гриле из нержавеющей стали или приготовленные путем погружения сырого мяса в открытое газовое пламя?

В данном случае более половины опрашиваемых выбрали метод «Макдоналдса», а когда исследователи указали, что жарки на открытом огне недостаточно и перед самой подачей гамбургеры в «Бургер-Кинге» разогреваются в микроволновой печи, процент за метод «Макдоналдса» вырос до 85. Мораль этой истории такова: лучшим ответом па результаты рекламного опроса является скептицизм. К сожалению, такие результаты подкрепляют распространенное мнение о том, что можно создавать и проводить опросы, которые будут давать практически любые желаемые результаты.

Мы привели лишь несколько примеров, иллюстрирующих трудности, которые возникают при разработке опросов. Существует несколько превосходных руководств по созданию надежных и валидных опросов (например, Converce & Presser, 1986; Fink, 1995; Flowler, 1993; Patten, 1998). Также обратите внимание на приложение *В* — в нем содержатся несколько полезных советов к разработке хорошего письменного опроса.

Четвертая и последняя проблема опросов не методологическая, а этическая и отчасти совпадает с первыми тремя. На основании результатов опросов иногда принимаются решения, способные повлиять на жизнь людей, и если эти опросы построены не вполне правильно, то люди могут пострадать. Хотя профессиональные психологи, опирающиеся на этический кодекс *АРА,* не станут использовать опросы для неподобающих целей. Тем не менее злоупотребления могут иметь место. Эта проблема не осталась без внимания законодателей, которые установили особые стандарты для использования данных опроса в суде (Morgan, 1990). В целом, данные стандарты говорят следующее: если вы собираетесь провести опрос и использовать полученные данные, подходите к этому так, как это делают профессиональные психологи, т. е. внимательно относитесь к выборке, структуре опросника, процедуре проведения опроса и анализу данных. Интересный, хотя и возмутительный пример злоупотребления приведен во вставке 12.2. В ней рассматривается случай ведущей программы новостей, которая была уволена с работы на основании данных некорректно проведенного опроса.

ВСТАВКА 12.2

Этика — опросы и злоупотребление ими

В начале 1981 г. телекомпания *KMBCв* Канзас Сити приняла журналистку из Калифорнии Кристин Крафт на должность ведущей программы вечерних новостей. Менее чем через год ее понизили в должности. Телекомпания заявила о ее некомпетентности. Мисс Крафт подала на компанию в суд, сказав, что ее нанимали на работу ради ее профессиональных способностей, а не ради привлекательной внешности. Она выиграла дело.

На первый взгляд казалось, что телекомпания повела себя весьма благоразумно и приняла свое решение не произвольно, а на основании данных, собранных в ходе опроса телезрителей. Однако данный пример - не иллюстрация использования метода опроса, а иллюстрация злоупотребления им. Исследование имело несколько важных дефектов (Beisecker, 1988). Во-первых, хотя опрос производил впечатление объективного, на самом деле он с самого начала проводился совершенно некорректно. Одна из основных проблем заключалась в том, что фирма, проводившая исследование, оказывала телекомпании платные консультационные услуги, и занимавшиеся опросом знали о том, какие результаты нужны руководству. Это отразилось на формулировании телефонных вопросов. Задача заключалась в сравнении мисс Крафт с другими ведущими местных программ новостей. Но вопросы были заранее составлены не в ее пользу. Например, Крафт показала низкий результат по вопросу «какой ведущий лучше знает Канзас Сити» (p. 23). И это неудивительно, если учесть то, что: а) на момент опроса мисс Крафт работала на телекомпанию *KMBC* всего 6 месяцев, а другие ведущие работали не меньше года и б) в начале ее работы в *KMBC* мисс Крафт усиленно рекламировали как «свежего человека из Калифорнии» (p. 23). Кроме того, очень мало вопросов было посвящено оценке журналистских способностей - большинство из них касались физической привлекательности (например: «Ее привлекательность - это большой плюс для программы новостей, которую она ведет» (p. 23).

Кроме внесенного таким образов искажения возникли достаточно серьезные проблемы со статистикой данных, позволившие Бейсекеру заключить, что «фирма, проводившая исследование, не использовала ни один известный метод статистической проверки или статистического обобщения, чтобы подтвердить выводы, сделанные на основании данных, полученных от конкретной выборки» (p. 25). Другими словами, консалтинговая фирма обобщила результаты опроса с помощью описательной статистики, но не воспользовалась методами заключительной статистики, позволяющими сделать выводы о значимости полученных результатов. Например, когда людей в возрасте от 25 до 34 лет спросили, на каком канале качество программы новостей снизилось, 30% ответили, что на *KMBC,* а 16% - что на конкурирующем канале *KCMO.* В отчете много говорилось об этом «соотношении 2:1». Его появление было отнесено на счет якобы пагубного влияния мисс Крафт. Но за год до этого (1980), когда мисс Крафт еще не работала в компании, та же фирма задала тот же вопрос в ходе другого опроса. В результате 26% людей указали КМВС и 15% - *KCMO.* Заключительный анализ несомненно не показал бы никакой разницы между данными этих двух лет. Анализ не просто не был проведен, консультанты/исследователи даже не включили данные 1980 г. в отчет 1981 г.

Очевидно, что опросы можно проводить правильно и можно корректно интерпретировать результаты. Но в случае с Кристин Крафт исследование не соответствовало стандарту 1.06 этического кодекса *APA* (American Psychological Association, 1992): «Психологи в своей научной или профессиональной деятельности и суждениях, опираются на научную информацию и профессиональные знания».

Дискуссии об использовании животных в качестве объектов исследования нередко скорее чадят, чем разгораются (см. главу 2). Один из способов узнать, что думают люди по этому поводу, — провести соответствующий опрос. В примере 33 рассказывается о таком опросе.

Пример 33. Опрос об исследованиях с участием животных

В 1996 г. Плаус опубликовал результаты двух опросов на тему исследований с участием животных. Первый из них был проведен среди психологов, занимающихся академической деятельностью (Pious, 1996a), а второй — среди студентов психологических специальностей (Pious, 1996b). Так как вы тоже входите в эту группу, рассмотрим последний случай более подробно.

В методологии опроса термин **база выборки** обычно означает полный список индивидуумов, из которого делается выборка. В случае групповой выборки база выборки означает список изначально отобранных групп. Плаус использовал именно такую процедуру. Начальная база выборки представляла собой список из 708 колледжей и университетов, выбранных в справочнике по вузам. В базу не вошли лишь несколько очень маленьких учебных заведений и религиозные вузы. Из 708 вузов с помощью случайной выборки отобрали 50, а затем связались с руководством психологических факультетов. 42 учебных заведения согласились участвовать в опросе. Затем провели выборки студентов каждого факультета. Конечные результаты были получены на основании 1188 анкет, возвращенных студентами психологических специальностей. Количество возвратов составило 58,75%. Довольно низкий показатель вызвал сомнения по поводу репрезентативности, и Плаус изучил выборку более подробно. Изучение правительственных данных за 1995 г. показало, что 73,1% студентов психологических специальностей — женщины, а 26,9% — мужчины. Это вполне соответствовало выборке, сделанной Плаусом: 73% женщин и 27 % мужчин.

Студентов предупредили о том, что опрос будет посвящен использованию животных в психологических, но не медицинских исследованиях (хотя некоторые совпадения возможны) и для образовательных нужд. Им задавали серии закрытых вопросов, в том числе те, которые приведены в табл. 12.3. Кроме того, студентов попросили заполнить таблицу из четырех столбцов, озаглавленных названиями видов животных (приматы, собаки, крысы и голуби), и трех рядов, различающихся по жестокости применяемых экспериментальных процедур (наблюдение; помещение в искусственные условия — процедура при этом пе вызывает боли и не приводит к смерти; физическая боль и смерть). Студенты должны были указать в каждой ячейке, будет ли подобное исследование «в обычном случае оправданно» или «в обычном случае неоправданно». Поскольку обсуждаемая проблема была весьма сложной, Плаус хотел, чтобы вопросы были ясными, понятными и не вносили искажений. Для этого была проведена предварительная проверка вопросов на другой выборке студентов психологических факультетов.

Представленные в табл. 12.3 результаты показывают, что студенты психологических факультетов одобряют использование животных и для исследовательских целей, и для целей обучения. Более 70% поддержали или активно поддержали исследования с участием животных, около двух третей сочли, что исследования с жи-

вотными необходимы для развития психологии, и хотя большинство ответило, что лабораторная работа с животными не должна входить в число обязательных занятий для студентов психологических специальностей, преобладающее число студентов поддержало использование животных в ходе занятий. Студенты с участием отнеслись к вопросам о причиняемых боли и страданиях и почти единодушно поддержали идею охраны психологического благополучия приматов, а также необходимость оценки уровня болевого воздействия, которое могут испытать животные. Для всех четырех видов, представленных в таблице 3x4, были получены ответы, демонстрирующие преобладание негативного отношения к процедурам, вызывающим боль и приводящим к смерти животных.

Таблица 12.3

Отношение к исследованиям с участием животных — опрос студентов психологических специальностей

В таблице представлена выборка из опроса, проведенного Плаусом (Pious, 1996b). Поддерживаете ли Вы в целом использование животных в психологических исследованиях?

Активно поддерживаю	14,3%
Поддерживаю	57,4%
Не поддерживаю	13,8%
Не поддерживаю категорически	4,7%
Затрудняюсь ответить	9,8%

Считаете ли Вы, что для развития психологии в психологических исследованиях необходимо использовать животных?

Да	68,4%
Нет	15,7%
Затрудняюсь ответить	15,9%

Поддерживаете ли Вы в целом использование животных при проведения в вузах занятий по психологии?

Поддерживаю	56,9%
Не поддерживаю	28,4%
Затрудняюсь ответить	14,7%

Как Вы думаете, должна ли входить работа с животными в число обязательных занятий на психологических факультетах?

Да	34,1%
Нет	54,3%
Затрудняюсь ответить	11,6%

Федеральные законы направлены на охрану «психологического благополучия» приматов, используемых в исследованиях. Поддерживаете ли Вы идею охраны психологического благополучия приматов?

Поддерживаю	85,1%
Не поддерживаю	4,0%
Затрудняюсь ответить	10,9%

Прежде чем получить разрешение па проведение исследовании, ученые в Великобритании, Канаде и Нидерландах должны оцепить уровень боли, который могут испытать животные. Поддерживаете ли Вы введение аналогичных требовании в Соединенных Штагах?

Поддерживаю	85,2%
Не поддерживаю	7,0%
Затрудняюсь ответить	7,8%

Источник: адаптировано из книги Плауса (Pious, 1996a), табл. 1.

Но каковы же были результаты другого опроса — посланного психологам, работающим в вузах? Интересно, что результаты двух опросов были практически одинаковыми. Единственное заметное различие было обнаружено в пункте 1. 71,1% студентов либо активно поддержали (14,3%), либо поддержали (57,4%) использование животных в психологических исследованиях. Но психологи оказали еще более сильную поддержку: 31,4% активно поддержали и 48,6% поддержали (т. е. 80% поддержали или активно поддержали) подобные исследования.

Исследование конкретных примеров

В начале предыдущей главы, где рассказывалось о том, что первые исследования по психологии проводились с малым *N,* я упомянул проведенный Дарвипом подробный анализ раннего детства своего сына. Такой анализ — это один из случаев **исследования конкретного примера,** процедуры, в широком смысле понимаемой как тщательный анализ отдельного случая. Обычно «случай» — это человек, а метод иногда называют «историей индивидуума», так как он подразумевает подробный анализ истории жизни конкретного человека или ее значительной части. Этот подход часто используется в клинической работе, когда случай пациента с определенным нарушением используется для иллюстрации факторов, приводящих к этому нарушению, или методов его лечения. Самый известный психолог, практикующий этот вид исследования конкретного примера, — Зигмунд Фрейд, создавший свою теорию личности на основании подробного изучения своих пациентов.

Истории индивидуумов составляются также психологами-экспериментаторами. Один из наиболее известных примеров — это потрясающее исследование Александра Лурии, посвященное человеку, неспособному забывать информацию. Это классическое исследование подробно описано во вставке 12.3, и если вы думаете, что совершенная память может избавить вас от многих проблем, то, прочитав о Ш., вы сможете по-другому взглянуть иа этот вопрос.

Исследования конкретных примеров не ограничиваются изучением отдельных людей. Этот термин используется и в случае анализа отдельного события или класса событий, отображающих определенное явление. Например, хотя проведенное Фестингером исследование религиозной общины было использовано во вставке 12.1 как известный пример наблюдения участников исследования, эту работу также можно охарактеризовать как исследование конкретного культа. Исследования конкретных событий нередко проводятся, когда происходит редкое или имеющее важное историческое значение событие.

' ВСТАВКА 12.3

Классические исследования — ум мнемониста

Истории индивидуумов нередко описывают жизнь ярких представителей отдельных психологических типов. Например, в анормальной психологии подробное изучение частных примеров нарушений используется, чтобы лучше понять динамику этих нарушений. Но метод составления истории индивидуума плодотворно используется и в экспериментальной психологии - он проливает свет на базовые психологические явления. Классическим примером метода служит работа Александра Лурии (1902-1977), советского психолога и физиолога, известного своим изучением мозговых травм раненых и исследованием взаимосвязи мышления и речи (Brennan, 1991).

Интересующий нас пример был посвящен изучению феномена С. В. Шерашевского, или Ш., чья исключительная память позволяла ему работать на сцене в качестве мнемониста (да, действительно, люди платили деньги, чтобы посмотреть, как он запоминает информацию) и причиняла ему множество неудобств. Этот пример описан в работе Лурии «Ум мнемониста», изданной в 1968 г. Лурия исследовал Ш. более 20 лет, фиксируя силу его памяти и проблемы, сопутствующие неспособности забывать. Лурия обнаружил отсутствие ограничений на объем информации, которую мог запомнить Ш., а самым удивительным было то, что информация не стиралась из памяти со временем. Ш. легко мог запомнить список из 70 чисел, а затем воспроизвести их в прямом и обратном порядке. Также «опыты показали, что он с успехом - и без заметного труда - может воспроизводить любой длинный ряд слов, данных ему неделю, месяц, год, много лет назад» (Luria, 1968, p. 12).

Способности Ш. невероятны, особенно если учитывать то, что большинство людей при выполнении подобных заданий не способны запомнить более семи-восьми элементов и что забывание - это скорее правило, чем исключение. Вас, как студентов, может заинтересовать обратная сторона этой способности: похоже, что неплохо было бы иметь такую «проблему», особенно во время экзаменов?

К сожалению, феноменальные возможности памяти Ш. сопровождались сильным дефицитом других познавательных функций. Например, он почти не способен был читать, понимая прочитанное. Это было связано с тем, что каждое слово вызывало в его памяти яркий зрительный образ, мешающий восприятию общего смысла, передаваемого фразой. Кроме того, Ш. плохо справлялся с решением задач, испытывал трудности с планированием и устройством своей жизни и был не способен к абстрактному мышлению. Сила визуальных образов, породившая удивительные способности его памяти, мешала протеканию других процессов. Является ли Ш. лишь любопытным, но бесполезным примером, единственным в своем роде странным человеком, чья история не поможет нам лучше понять самих себя? Конечно, нет. Данный случай проливает свет и на работу обычной памяти, в частности на функциональный смысл ограничения возможностей краткосрочной памяти. Иногда мы ругаем себя за неспособность запомнить что-то, о чем мы думали лишь несколько минут назад, но случай Ш. показывает, что забывание позволяет нам очистить сознание от бесполезной информации (например, ни к чему запоминать все встречающиеся нам номера телефонов) и помогает сконцентрировать энергию на более сложных заданиях, например на чтении. Поскольку Ш. не мог не запоминать, он был не способен к работе на более высоких познавательных уровнях.

В заключение отметим, что Ш. был не единственным случаем подобного рода. Американские психологи Хант и Лав (Hunt & Love, 1972) исследовали другого человека (ВП) с феноменальной памятью. Удивительно, что ВП вырос в латвийском городе, расположенном неподалеку от родного города Ш.

Например, ученые исследовали, как люди, живущие вблизи от атомной станции Три Майл Аплэнд в штате Пенсильвания, отреагировали на произошедшую там в 1979 г. аварию (Aronson, 1999a) и насколько точные «моментальные» воспоминания сохранились у людей о том, где они были и что они делали в момент убийства президента Кеннеди в 1963 г. (Yarmen & Bull, 1978).

Пример 34. Гипноз на сцене

Кроме изучения последствий редких и необычных событий исследование конкретных примеров может быть посвящено событиям весьма распространенным, но почему-либо интересным. Например, это может быть публичная демонстрация гипнотизма, в ходе которой обычные люди делают весьма необычные вещи. Именно этому явлению было посвящено исследование Кроуфорда с соавторами (Crawford, Kitner-Triolo, Clarke, & Olesko, 1992). Исследователи провели опрос 22 студентов колледжа, которые участвовали в сеансе гипноза, проведенном в их университетском городке. Задача исследования состояла в изучении последствий подобного переживания. Для большинства студентов состояние транса оказалось весьма приятным, но некоторые нашли его беспокоящим, унизительным или даже пугающим. Удивительным было то, что значительная часть группы испытала состояние постгипнотической амнезии. Восемь из 22 студентов согласились со следующим утверждением: «Покинув сцепу, я не мог вспомнить некоторые вещи, которые я помнил раньше». Один участник, особенно сильно обеспокоенный амнезией, сравнил свое состояние с отключением сознания, вызванным алкоголем. Также необычным было и то, что 5 из 22 студентов утверждали, что гипнотизер абсолютно подчинил себе их действия. В лабораторных исследованиях гипноза обычно не обнаруживается подобная степень амнезии и утраты воли. Таким образом, данное исследование конкретного примера дало важную информацию о таком воздействии гипноза, которое не было обнаружено в ходе исследований со специально созданными условиями.

На основании проведенного анализа Кроуфорд и др. разработали серию рекомендаций к проведению публичных сеансов гипноза, в том числе: «устранять от участия людей, проходящих лечение, создавать у участников правильное представление о гипнозе до его применения... удалять любые признаки амнезии, отслеживать события, происходящие при выходе участника из транса, и обеспечивать участникам возможность после сеанса задавать возникшие у них вопросы» (р. 666).

Кроме этого, термин «исследование конкретного примера» или просто «пример» может применяться для обозначения хорошего примера какого-либо явления. Именно в этом смысле я использовал его в дайной книге: начиная с главы 4 и до только что рассмотренного случая гипнотизма вы встретились с различными «примерами», представляющими собой весьма подробные описания исследований, выбранных потому, что они хорошо иллюстрируют определенные методы или экспериментальные планы.

Оценка исследования конкретных примеров

На первый взгляд может показаться, что описание отдельных случаев не слишком информативно. Самое явное слабое место таких исследований — это внешняя валидность: как можно делать обобщения на основании лишь одного примера? Конечно, для результатов любых исследований уверенность во внешней валидности

нарастает постепенно, по мере воспроизведения сделанных открытий. Это верно и для исследования конкретных примеров. Воспроизвести такое исключительное событие, как землетрясение в Сан-Франциско 1989 г., может быть трудно, но можно исследовать, проявляется ли аналогичное поведение при других стихийных бедствиях.

Вторая трудность в использовании конкретными примерами заключается в том, что у исследователей есть богатые возможности для того, чтобы придать описанию желаемый оттенок и тем самым исказить исследование. Так ли удивительны бесконечные дискуссии по поводу эдипова комплекса и других неразрешимых сексуальных проблем, встречающихся в историях индивидуумов, написанных фрейдистами? Занимаясь поисками примеров для этой главы в базе данных *PsycINFO,* я встретил исследование под названием «Зависть к докторской степени: психоаналитическое исследование конкретного примера» (Behr, 1992). В нем рассказывалась история женщины с «глубокой выматывающей тревожностью в связи с невозможностью закончить докторскую диссертацию... [Ее] презрение к тем, кто не имел докторской степени, и сильная зависть к ее обладателям сделала практически равноценными зависть к докторской степени и зависть к пенису» (р. 99). Разве не иначе описал бы этот случай противник фрейдизма? Наверняка иначе.

Третья проблема касается памяти. Поскольку при составлении историй конкретных индивидуумов участникам часто требуется вспоминать события прошлого, возникает ряд проблем, которые связаны с несовершенством памяти. Исследования, проведенные Элизабет Лофтус и другими учеными, неоднократно показали, что воспоминания о событиях нашей жизни нередко искажаются другими событиями, произошедшими между интересующим случаем и попыткой его позднейшего припоминания. Возьмем, к примеру, жителя Флориды, Эндрю, пережившего в 1992 г. чрезвычайно сильный ураган. Если через год в ходе исследования конкретного примера его попросить описать произошедшее, то некоторая информация без сомнения окажется точной. Ведь маловероятно, что человек забудет, как сорвало крышу с его дома. Но за этот год человек: а) пережил само событие, б) просмотрел видеофильмы, новости, телевизионный показ и фотографии данного события, в) прослушал бесчисленные истории об урагане, рассказанные друзьями и соседями, и г) возможно, даже несколько раз видел сны о данном событии. Как выявили Лофтус и ее исследовательская группа (например, Loftns & Hoffman, 1989), более поздние воспоминания часто представляют собой «построения» из впечатлений от самого события и от явлений, имевших место впоследствии. Таким образом, мы имеем воспоминание, включающее множество элементов, и само событие — только один из них.

Несмотря на эти трудности, истории индивидуумов могут быть весьма информативными. Во-первых, благодаря большому количеству подробностей они позволяют глубже понять поведение, чем это возможно при исследованиях с большим *N,* при которых от каждого испытуемого получают весьма ограниченный объем информации. В исследовании Лурип были не только подробно описаны возможности уникальной памяти Ш., но и показано, как его память влияла на другие аспекты жизни несчастного человека. Во-вторых, случай Ш. иллюстрирует также и другой момент: некоторые индивидуальности встречаются настолько редко, что единственно возможным исследованием является исследование конкретного примера.

В-третьих, как и все описательные исследования, изучение конкретных примеров может стать богатым источником эмпирических вопросов для дальнейших изысканий. Как было отмечено выше, Кроуфорд и его коллеги (Crawford и. др., 1992) были удивлены тем, что около трети студентов колледжа, участвовавших в сеансе демонстрации гипноза, испытали амнезию и в той или иной степени забыли происходившее на сцене, а также тем, что почти все студенты считали, что гипнотизер полностью контролировал их действия. Повлияли ли на это какие-либо когнитивные или личностные факторы, отличающие этих студентов от остальных? В чем сходны и чем различаются постгипнотическая и посталкогольная амнезии? Будет ли вероятность развития приобретенной беспомощности у студентов, сохранивших память о полном контроле, выше, чем у тех, кто чувствовал себя более уверенно?

Исследование конкретных примеров не только рождает гипотезы дальнейших исследований, но и служит целям фальсификации. Если о некотором психологическом явлении высказывается определенное утверждение, то даже частное опровержение может вызвать серьезные сомнения в этом утверждении. Вы уже познакомились с одной известной иллюстрацией этого (вставка 3.3). Утверждение о наличии у Умного Ганса математических способностей было успешно фальсифицировано, после того как этот случай был исследован психологом Пфунгстом.

Второй пример взят из области парапсихологии, исследующей, в частности, такое явление, как экстрасенсорное восприятие. Вы наверняка читали о «медиумах», которые утверждают, что умеют читать мысли, предсказывать будущее и действием психической силы вызывать определенные события (например, сгибать ложки). Подобные утверждения можно поставить под вопрос, указав на исследование конкретного примера, в котором человек (обычно профессиональный фокусник) воспроизводит все продемонстрированные медиумом явления, но показывает, что все это было проделано с помощью обычной ловкости рук (например, Randi, 1982). Защитники парапсихологии могут возразить, что медиумы добиваются результата психическими средствами, а фокусники — физическими, но более простое объяснение заключается в том, что и те и другие пользуются одними и теми же средствами (т. е. физическими).

В целом, исследования конкретных примеров подвержены искажениям и их результаты, может быть, нелегко обобщить, но, как и наблюдения, они полезны для разработки новых исследований, помогают фальсифицировать слабые теории и, кроме того, иногда являются единственным способом документирования информации о необычном человеке или событии.

Архивное исследование

Иногда для проверки гипотез исследователи используют ранее полученную информацию. Такой подход называется **архивным исследованием,** а термин «архив» означает как сами записи, так и помещение для их хранения. Архивные данные включают весь спектр информации от общедоступной, такой как данные переписи населения, судебные протоколы, генеалогические сведения, годовые отчеты фирм и регистрационные данные из бюро патентов, до более персональной, как, например, кредитные отчеты, истории болезней, отметки образовательных учреждений,

личная корреспонденция и дневники. «Архивами» в смысле помещений могут быть университетские библиотеки, правительственные учреждения или компьютерные базы данных. Прочитав приведенный ниже список названий статей, вы получите представление о вопросах, изучаемых с помощью архивных исследований:

«Восприятие справедливости, мотивация и арбитраж в высшей бейсбольном лиге» (Bretz & Tomas, 1992);

«Узаконивать гражданского брака: оценка, проведенная на основании документов о разрешении на вступление в брак от 1970 и 1980 гг.» (Gwaltney-Gibbs, 1986);

«Авторство в колонках советов: действующие лица и наблюдатели, причины и поводы» (Schoeneman & Rubanowitz, 1985)

Виды архивных исследований

Архивные исследования часто являются частью более крупных исследовательских проектов, проводимых с применением различных методов. Показательный пример работы такого рода — исследование по оценке программ. Как вы помните из главы 10, исследования по оценке программ включают анализ потребности, предварительную и суммарную оценки, а также анализ затрат. Все эти процедуры могут проводиться на основе архивной информации. При проведении анализа потребности обычно изучаются данные переписи населения, а проведению анализа стоимости и обоих видов оценок могут помочь данные психологических организаций и статистическая информация, полученная при проведении других программ. Кроме того, обычно архивная информация (например, данные об автомобильных авариях) входит в один из рассмотренных в главе 10 методов — план с прерванной временной последовательностью, — и изучается для выявления тенденций.

Иногда архивные исследования требуют не только отбора и статистического анализа записей, но также контент-анализа информации. Контент-анализ — это любое систематическое изучение материала, в ходе которого качественная информация распределяется по заранее определенным категориям. Любой вид записей можно подвергнуть контент-анализу; и хотя в большинстве случаев изучается вербальная информация, например данные интервью, невербальный материал также можно проанализировать. Например, в исследовании Брэптжета и Боума (Brantjets & Bouma, 1991) рисунки пациентов, страдающих болезнью Альцгеймера, подвергли контент-анализу, затем установили соотношение рисунков со степенью психической деградации и сравнили их с рисунками людей того же возраста, не страдающих этим недугом.

При проведении наблюдений заранее разрабатываются контрольные таблицы. В данном случае важно операционально определить категории материала, подвергающегося контент-анализу. Также неплохо, чтобы анализ проводили несколько исследователей — это позволит удостовериться в его надежности. Чтобы избежать искажений, аналитики по возможности не должны знать, в чем состоит проверяемая гипотеза.

Исследование Ло и Рассела (Lau & Russel, 1980) показывает, как можно эффективно использовать контент-анализ вербального материала. Чтобы узнать, как люди объясняют свои спортивные удачи и поражения, исследователи изучили газетные статьи о крупных спортивных событиях, содержащие комментарии игро-

ков, тренеров и спортивных обозревателей о том, почему они (или команда, о которой они пишут) выиграли или проиграли. Полученные объяснения были разделены на две большие категории в зависимости от того, считались ли победа пли поражение вызванными внутренними (целиком наша вина) или внешними (какого-либо рода оправдание) факторами. Исследователи обнаружили, что победители обычно дают «внутреннее» объяснение своему успеху, а у проигравших наблюдается примерно равное количество «внутренних» и «внешних» объяснений.

Хотя по своей природе архивные исследования не позволяют непосредственно управлять независимой переменной, исследователи нередко пытаются весьма непростыми способами контролировать осложняющие факторы. Довольно неглупый этому пример — исследование Ульрнха (Ulrich, 1984), в ходе которого было показано, что на выздоровление после хирургических операций может повлиять наличие в больничной палате приятного вида из окна.

Пример 35. Комната с видом

Одна из наиболее интересных областей исследований в психологии, разрабатываемая в последние 30 лет, касается взаимосвязи психического и физического здоровья людей. Психологи, исследующие вопросы здоровья, изучают такие проблемы, как связь между стрессом и болезнью, взаимоотношения врачей и пациентов и, как в исследовании Ульрнха, влияние такого на первый взгляд незначительного фактора, как архитектура больничных помещений на здоровье пациентов.

Ульрих изучил истории болезней, составленные за 10 лет в пригородной пенсильванской больнице. Он хотел узнать, влияют ли на процесс выздоровления особенности помещения, в котором пациенты проводят послеоперационный период. В частности, он сравнил помещения, в которых окна выходили на кирпичную стену, с другими, с видом из окон на небольшую группу деревьев (чертеж здания представлен на рис. 12.2). Все палаты были одинакового размера и кровати в них стояли так, чтобы больные легко могли посмотреть в окно.

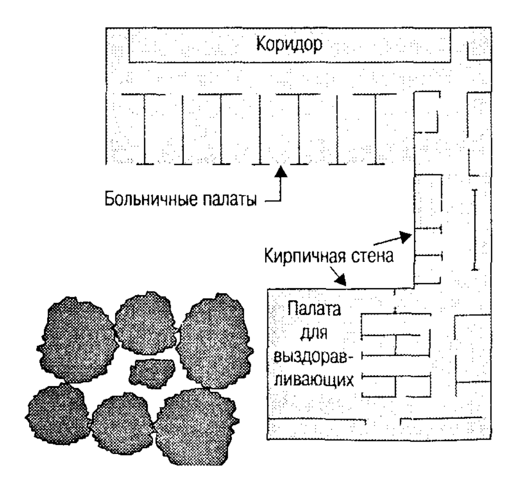


Рис. 12.2. Схема здания больницы, которую использовали в архивном исследовании восстановительного периода после хирургических операций. На схеме показаны палаты с видом на группу деревьев и на кирпичную стену

Поскольку исследование было архивным, Ульрих не мог случайным образом распределить пациентов по палатам, но сделал все возможное, чтобы две группы больных можно было сравнивать. Во-первых он учитывал только пациентов, выздоравливающих после довольно часто встречающейся операции на желчном пузыре, и не рассматривал больных моложе 20 и старше 69 лет. Во-вторых, использовав различные переменные уравнивания, в том числе возраст, пол, курение и вес, он создал две схожие группы. Таким образом, он изучил истории болезней у двух приблизительно эквивалентных групп пациентов — единственным существенным различием между ними был вид из окна в их палатах: кирпичная стена или деревья.

По архивным данным Ульрих изучил сроки послеоперационного лечения, установил, получали ли больные, и если получали, то в каком количестве, обезболивающие и успокоительные средства, фиксировал частоту жалоб на головную боль или тошноту, а также записи о пациентах медицинских сестер. Ученый обнаружил явное преимущество выздоравливания в палатах с хорошим видом. Пациенты, которые лежали в таких палатах, проводили в больнице после операции в среднем на один день меньше (7,96 и 8,70), им требовалось меньше обезболивающих, а если они употребляли их, то в меньших дозах, чем пациенты из палат с видом на стену. Контент-анализ записей медицинских сестер показал, что к больным из палат с видом на деревья чаще относились положительные, а не отрицательные отзывы (но, возможно, сестры также находились под впечатлением от вида на деревья!). Группы не различались ни по применению успокоительных средств, ни по жалобам на послеоперационное недомогание.

Исследование Ульрнха, как и большое число других исследований, демонстрирует, как многочисленные внемедицинские факторы влияют на физическое здоровье. Также оно показывает, благодаря чему психология здоровья стала междисциплинарной наукой: к моменту публикации исследования Ульрих преподавал на факультете *географии* в Делавэрском университете.

Оценка архивных исследований

Архивные исследования имеют несколько преимуществ по сравнению с другими видами исследований. Первое из них состоит в том, что сегодня, в начале XXI в., доступен необъятный объем информации и возможности проведения архивных исследований ограничены лишь креативностью исследователя. Вторая сильная сторона архивных исследований заключается в том, что они могут согласовываться с результатами лабораторных исследований, тем самым увеличивая внешнюю валидность. Пример тому — исследование Мак-Клелландом мотивации достижения, о котором вы прочитали во вставке 9.3. В исследовании была использована архивная информация: чтобы показать, что книжные истории о достигнутом успехе влияют на будущие достижения читателей, изучались детская литература и информация о патентах, полученных в разные исторические моменты. Результаты вполне согласовывались с данными лабораторных исследований, посвященных изучению взаимосвязи между стремлением индивидуумов к успеху и поведением, направленным на достижение успеха. Третье достоинство этого метода состоит в том, что архивная информация не подвержена действию реактивности: она уже зафиксирована.

Тот факт, что архивная информация уже зафиксирована, создает и проблемы. Несмотря на объем доступных данных, какая-то важная информация может отсутствовать, а другие данные могут не обладать репрезентативностью. Например, в одном из упомянутых ранее исследований газетных колонок советов Шенеман и Рубановиц (Schoeneman & Rubanowitz, 1985) столкнулись с тем, что публикуемая в таких колонках информация представляет собой лишь малую часть писем, получаемых составителями этих рубрик. Кто может сказать, какие факторы определяют отбор писем для публикации?

Последняя проблема архивных исследований, связанная со всеми описанными в данной главе методами, — искажение, вызванное экспериментатором. При проведении архивного исследования легко можно ошибиться и отобрать только ту информацию, которая подтверждает выдвинутую гипотезу, а также интерпретировать данные сквозь призму определенных ожиданий. Эту проблему бывает трудно решить полностью, так как проводящий архивную работу исследователь обычно находит гораздо больше информации, чем может использовать, и эту информацию нередко можно различно интерпретировать.

Конечно, рассудительный исследователь может снизить влияние искажений, заранее разработав четкий план, тщательно продумав процедуру исследования, операциональные определения всех имеющих отношение к делу переменных, критерии отбора материала и т. д. Например, в исследовании больничного выздоровления изучались только пациенты, поступившие в больницу между 1 мая и 20 октября рассматриваемого года. Почему были выбраны именно эти месяцы? Потому что в остальные месяцы года на деревьях нет листвы, а Ульрих хотел сравнить пациентов, наблюдающих кирпичную стену, и больных, имеющих возможность любоваться зелеными деревьями.

Таким образом, архивные исследования являются эффективным способом использования информации об имевших место в прошлом событиях и тем самым помогают ученым проверять гипотезы при отсутствии возможности непосредственного управления переменными. Учитывая огромное количество архивной информации и ставший более легким доступ к ней (например, через Интернет), можно ожидать, что в будущем архивные исследования станут еще более популярными.

Эта глава, посвященная описательным методам, завершает работу под названием «Исследование в психологии». Остается лишь краткий эпилог, который начинается с обращения к рассмотренной в главе 1 теме удовольствия, получаемого от проведения исследований по психологии, и заканчивается рассуждениями о том, какое влияние мог оказать на вас курс методологии исследований.

Резюме

Наблюдение

С помощью основных методов наблюдения можно дать правильное качественное и количественное описание поведения. При естественном наблюдении наблюдатель отделен от тех, кого наблюдает, а субъекты исследования либо не знают о присутствии наблюдателя, либо привыкают к нему. При наблюдении участников исследования наблюдатель становится активным членом изучаемой группы. Наблю-

дения различаются по степени воздействия экспериментатора на наблюдаемые условия. Методология наблюдения учитывает три проблемы: отсутствие контроля, искажение, вносимое наблюдателем, и реактивность субъектов. Кроме того, исследователи сталкиваются с этическими проблемами: проблемой получения обоснованного согласия и проблемой вторжения в личную жизнь. Так же как и другие виды описательных исследований, рассмотренные в этой главе, наблюдения могут использоваться для фальсификации утверждении о поведении и нередко становятся источником гипотез дальнейших изысканий.

Опрос

Главная цель опроса состоит в сборе описательной информации о мнениях людей, их отношении к разным вещам и об их представлении о собственном поведении. Опросы могут проходить в виде персональных интервью, в письменной форме или по телефону. Разработка надежного и валидного опроса — это важная задача, ведь на данные опроса могут повлиять искажения, вносимые неполным возвратом опросников и ожиданиями опрашивающих и опрашиваемых. Если отобранная группа не отооражает популяцию, на которую нацелен опрос, то сделанные на основании опроса выводы могут быть неверными.

Исследование конкретных примеров

Исследование конкретного примера дает подробное аналитическое описание отдельного человека или события. Истории индивидуумов составляются для описания прошлого и характерных особенностей людей, входящих в определенную категорию (например, индивидуумов, имеющих определенную психологическую проблему, необычных людей). Изучение конкретных событий проводится, если исследователи имеют возможность собрать данные о каком-либо значительном событии (например, землетрясении) или исследовать какое-то представляющее интерес явление (например, публичный сеанс гипноза).

Архивное исследование

Если исследователь использует информацию, которая собрана с какой-либо другой целью, или входит в общедоступную базу данных, то такое исследование называется архивным. Анализ архивной информации нередко является важной частью исследования по оценке программы, особенно при использовании плана с прерванной временной последовательностью. Архивные исследования могут разрабатываться, чтобы получить ответы на эмпирические вопросы, но их результаты ограничены пределами доступной информации, и поэтому нет никакой гарантии, что данные будут репрезентативными.

Задания для повторения

Выбор ответа

1. Для какого вида исследований наиболее вероятно установление таких сильных взаимосвязей экспериментатора с группой, что он теряет объективность?
   1) наблюдение религиозной общины участниками исследования;

2) групповой опрос, проходящий в виде интервью;
3) естественное наблюдение группы приматов;
4) исследование конкретной группы социальных работников.

2. Как решается проблема реактивности при проведении естественных наблюдений?
   1) субъекты не знают, что за ними наблюдают;
   2) субъекты знают, что за ними наблюдают, и дают обоснованное согласие;
   3) наблюдатели скрыты или субъекты привыкают к их присутствию;
   4) наблюдатель является членом изучаемой группы.
3. Основное преимущество метода интервью перед письменными и телефонными опросами заключается в том, что интервью:
   1) дают более полную и глубокую информацию;
   2) требуют меньших затрат на проведение;
   3) не порождают проблемы репрезентативной выборки;
   4) не вносят искажений.
4. На какой из перечисленных ниже эмпирических вопросов **лучший ответ** даст архивное исследование?
   1) Какой вид тропических рыб активнее всего охраняет свою территорию?
   2) Как большинство американцев относятся к идее государственного страхования здоровья?
   3) Действительно ли стихийные бедствия активнее освещаются средствами массовой информации, если они происходят в местах, часто посещаемых туристами?
   4) Действительно ли в кинотеатрах зрители с лишним весом чаще покупают «ведро» попкорна, а зрители с нормальным весом — «стакан»?
5. Какое из следующих утверждений об исследовании конкретных примеров как о методе верно?
   1) исследования конкретных примеров всегда включают подробное изучение отдельных индивидуумов;
   2) по сравнению с рассмотренными в данной главе методами исследование конкретных примеров позволяет наиболее верно установить причинно-следственные связи;
   3) этот метод может быть единственно правильным выбором при изучении индивидуумов с редкими особенностями;
   4) возникновение искажения, внесенного экспериментатором, маловероятно.

Короткие эссе

1. Чем с методологической и этической точек зрения отличаются процедуры проведения естественных наблюдений и наблюдений участников исследования?
2. В чем состоит проблема «реактивности» и как ее решают?

3. Покажите на примере, как искажение, внесенное наблюдателем, может повлиять на наблюдение и опишите, как можно снизить это искажение.
4. Изложите достоинства и недостатки интервью как метода проведения опроса.
5. Что вы можете сделать для увеличения процента возврата при проведении письменного опроса? Что такое искажение, вызванное отсутствием ответа?
6. На примере рекламы гамбургеров покажите опасность искажения, вызываемого формулировкой вопросов.
7. На примере случая журналистки Кристин Крафт покажите, как возможно злоупотребление данными опроса.
8. В чем заключаются достоинства и недостатки проведения опроса по телефону.
9. Каковы две разновидности исследований конкретных примеров и в чем состоят сильные и слабые стороны этого метода?
10. Какая независимая переменная была использована в архивном исследовании Ульриха (вид из окна палаты) и какие переменные контролировались, чтобы создать сравнимые группы?
11. Изложите достоинства и недостатки архивных исследований. Какие из описанных в предыдущих главах методы могут использовать архивные данные?

Упражнения

Упражнение 12.1. Усовершенствование опросников

Ниже приведены несколько закрытых вопросов и утверждений, предполагающих ответ типа согласен/не согласен, которые могут использоваться в опросах. Для каждого из них: а) определите, в чем состоит ошибка формулирования, и б) перепишите вопрос или утверждение, исправив ошибку. В выполнении этого задания вам поможет приложение *В*.

1. Страдали ли Вы расстройством желудка в последнее время?
2. Люди с большим числом предубеждений обычно недружелюбны и не слишком умны.
3. насколько молод, по Вашему мнению, средний курильщик?
4. Согласны ли Вы с большинством в том, что нарушение закона о ремнях безопасности должно строго наказываться?
5. Большинство врачей высокомерны.
6. Люди с лишним весом не энергичны и обычно несчастны.

Упражнение 12.2. Построение определений поведенческих актов для исследований, проводимых с использованием метода наблюдения

Представьте себе, что вы начинаете исследование с использованием метода наблюдения и должны разработать четкие операциональные определения видов поведения, которые будете наблюдать. Для каждой из следующих гипотез дайте операциональное определение поведенческих актов, которые будут фиксироваться,

и опишите, как бы вы провели исследование (т. е. где вы будете его проводить, будете ли вы определенным образом воздействовать на наблюдаемые условия и т. д.).

1. Мужчины при разговоре перебивают женщин чаще, чем женщины мужчин.
2. При самостоятельной игре более старшие дети объединяются в группы, а более младшие предпочитают играть поодиночке.
3. Большинство студентов колледжа по-настоящему занимаются менее половины того времени которое , проводят в библиотеке.
4. Подростки, обходящие торговый центр, следуют строго определенным маршрутам, и повторяют свои походы через определенные промежутки времени.
5. Пары обмениваются взглядами чаще до замужества, чем после него.
6. Собаки более агрессивны на своей территории, чем за ее пределами.

Упражнение 12.3. Выбор описательного метода

Для каждой из приведенных ниже гипотез выберите наилучший методологический подход. Объясните ваш выбор. Для некоторых заданий возможен более чем один правильный ответ.

1. Женские и мужские клубы и объединения представляют собой регламентированные организации, вознаграждающие выполнение неписаных норм поведения.
2. Когда опрашиваемых просят объяснить их мнение по поводу спорных вопросов, они обычно опираются на единичные свидетельства.
3. Для опроса потребителей на тему о любимых закусках использована общенациональная выборка (2000 человек).
4. В течение одного месяца после получившего большую огласку самоубийства наблюдалось увеличение количества аварий со смертельным исходом, в которых водитель находился в машине один и в которой другие автомобили не участвовали.
5. После тяжелой аварийной посадки, при которой отказал механизм приземления и людей пришлось эвакуировать, большинство пассажиров рейса 304 пережили состояние нарушающей сон тревожности средней и высокой степени тяжести.
6. Посещая кафе, студенты колледжей чаще студенток появляются без компании, особенно во время обеда.
7. Нужно проверить результат анализа объявлений в газетах, который подтвердил гипотезу эволюционной психологии о том, что мужчины предпочитают более молодых (т. е. более фертильных) женщин, а женщины — более взрослых (т. е . тверже стоящих на ногах, способных содержать семью) мужчин.
8. При изменении содержимого кормушки для птиц соотношение посещений кормушки двумя различными видами птиц изменится за три дня.

Эпилог

Я искренне надеюсь, что вам понравился данный курс но исследовательским методам и ваши занятия были вполне успешными. Данный курс может быть трудным и иногда утомительным, но я уверен в том, что этот предмет — один из важнейших в учебном плане (история психологии на втором месте). В рекомендательных письмах, которые я пишу для своих студентов, поступающих в аспирантуру или ищущих работу, я всегда обращаю внимание на успехи по моему предмету и указываю навыки, которые были приобретены по мере изучения материала. Я бы хотел закончить книгу: а) перечислив эти навыки (о некоторых из них шла речь в главе 1) и б) вернувшись к рассмотренному в главе 1 вопросу о страсти исследователей к своей работе.

Что я узнал о методах исследований

Чтобы стать профессиональным психологом, нужно получить ученую степень (предпочтительно докторскую), а поскольку учеба в аспирантуре предполагает большое количество проводимых исследований, курс исследовательских методов является первым шагом на пути приобретения необходимых для этого навыков. Но большинство студентов психологических факультетов не собираются получать докторскую степень по психологии, и студенты иногда спрашивают, насколько им нужен данный курс, если они хотят после окончания колледжа начать работать. Если у вас такие же планы, я хочу заверить вас, что шансов получить достойную работу и продвигаться по службе будет больше, если вы успешно пройдете данный курс. Ниже приведен список полезных для будущей карьеры навыков, которые начнут формироваться, когда вы будете изучать этот предмет. Естественно, если вы идете в аспирантуру, эти навыки также будут вам полезны и увеличат ваши шансы на успех при решении задач, которые возникнут в ходе диссертационного исследования.

- Способность к эмпирическому мышлению.
  - Умение формулировать вопросы, ответ на которые должны дать результаты исследования.
  - Умение делать основанные на данных выводы о поведении.
  - Понимание того, что основанные на данных выводы представляют собой «рабочие истины», требующие новых исследований и повторения уже сделанных.
  - Умение логически мыслить в ходе исследования (как индуктивно, так и дедуктивно).
- Способность к критическому изучению информации о поведении.
  - Скептическое отношение к необоснованным, излишне сложным и/или простым утверждениям.
  - Умение строить сильные (т. е. причинно-следственные) выводы, основанные только на корреляционных данных.

  - Умение внимательно изучать графики и распознать вводящую в заблуждение информацию.
  - Умение дать альтернативные, более простые объяснения данным, на основе которых делаются определенные утверждения.
  - Скептическое отношение к утверждениям, основанным на сомнительных свидетельствах (например, на единичных данных, на отзывах клиентов, на данных наблюдений с внесенным искажением, на данных недостаточного количества наблюдений).
  - Недоверие к большинству утверждений о фактах поведения как о чем-то «доказанном» (как в высказывании: «Доказано, что прослушивание музыки Моцарта увеличивает *IQ* ваших детей»).
- Способность читать и осознанно воспринимать сложную информацию.
  - Умение читать учебники и статьи об исследованиях.
- Способность видеть скрытые проблемы, формулировать их и разрабатывать план их решения.
  - Опыт разработки и реализации эмпирических исследовательских проектов.
- Способность к эффективному поиску важной информации и к ее логической организации.
  - Опыт работы с *PsycINFO* и другими электронными средствами поиска.
- Способность к упорядочиванию, обобщению и интерпретации данных.
  - Умение написать лабораторный отчет и обобщить данные, приобретенные в упражнениях на сбор данных.
  - Умение применить методы описательной статистики и статистики вывода для оценки собранных в лабораториях данных.
- Усовершенствованные навыки составления отчетов.
  - Правила написания лабораторных отчетов в формате *APA* помогают быть точными в своих высказываниях и использовать наиболее простые формулировки.
- Усовершенствованные навыки общения.
  - Опыт устных презентаций, стендовых докладов и лабораторных отчетов.
- Усовершенствованные компьютерные навыки.
  - В особенности навыки работы со статистическими пакетами программ (например, *SPSS),* но также и работы с программами для подготовки презентаций (например, *PowerPoint)* и, возможно, с пакетами программ или веб-сайтами, адресованными студентам, планирующим эксперимент.

Список впечатляет. Но обратите внимание на мои слова о том, что эти навыки лишь *начнут* у вас формироваться. Для их дальнейшего совершенствования (если вы продолжите обучение на психологическом факультете и сохраните интерес к исследованиям) я могу порекомендовать вам следующее.

- Ищите возможность проведения независимых исследований в рамках проектов ваших преподавателей. Они будут рады вашей помощи в лаборатории,

и хотя сначала вам могут поручить выполнение рутинных заданий (например, ввод данных), очень скоро вы начнете принимать более полное участие в проектах. Возможно, что к моменту окончания учебного заведения вы даже проведете как ведущий исследователь одно или два изыскания, которые отразят ваши идеи.

- Проведите независимое исследование или выполните дипломную работу, если ваш вуз предоставляет такую возможность.
- Выбирайте факультативные предметы, предполагающие лабораторные занятия. Обычно это курсы когнитивной психологии, физиологической психологии и изучения ощущения/восприятия.
- Посещайте исследовательские конференции. Представьте публике ваш проект (лучше изобразить его в виде плаката). Вам необходимо посетить по меньшей мере одну-две студенческие исследовательские конференции. Эти небольшие конференции специально созданы, чтобы студенты имели возможность рассказать о своих работах. Обычно они длятся всего один день и состоят из стендовых докладов, нескольких устных докладов и выступления видного исследователя. Но лучше отправьте свою работу в оргкомитет региональной психологической конференции. Такие конференции обычно проходят в течение трех дней. Во время конференции можно присутствовать при интересных дискуссиях известных психологов и пообщаться с другими студентами-исследователями. Наконец, обдумайте посещение одной из национальных конференций, либо Американской психологической ассоциации (ежегодно в августе), либо Американского психологического общества (ежегодно в июне). Они представляют собой расширенный аналог региональных конференций. Обратите внимание на сроки. Чтобы представить исследование на региональной или национальной конференции, вам потребуется отправить на рассмотрение ваши работу или тезисы к докладу где-то в ноябре или декабре. Поэтому лучше начинать исследование не позже середины предпоследнего курса.
- Постарайтесь вступить в «Пси Хи» — национальное почетное психологическое общество (www.psichi.org). Кроме дополнительного плюса для вашего CV вы получите дополнительные возможности для проведения исследований. «Пси Хи» выдает несколько премий за исследования, кроме того, «Пси Хи» спонсирует стендовое представление исследований членов этой организации на всех перечисленных региональных и национальных собраниях.
- Отправьте вашу работу в журнал, специализирующийся на студенческих исследованиях. Наиболее известное такое издание — это журнал, финансируемый «Пси Хи», но есть и другие. Их список с адресами веб-сайтов можно найти на сайте «Пси Хи».
- Выберите интересующую вас тему и сосредоточьтесь на ней — у вас начнет формироваться страсть к исследованиям.

Таким образом, мы подошли к последнему разделу книги.

Страсть к психологическим исследованиям (часть 2)

Незадолго до конца долгой и плодотворной жизни известного русского психолога Ивана Павлова группа студентов попросила ученого написать небольшую статью для их журнала о том, что необходимо, чтобы стать великим психологом. Павлов написал, что нужны три вещи — систематическое обогащение своих знаний, скромность и страсть. Павлов писал: «Большого напряжения и великой страсти требует наука от человека. Будьте страстны в ваших исканиях!» (Babkin, 1949, p. 110)

Из главы 1 вы вкратце узнали о двух известных ученых-психологах: Элеанор Гибсон и Б. Ф. Скиннере, и узнали о радости, которую приносят совершаемые в ходе психологических исследований открытия. Я постарался подчеркнуть, что наилучшая причина заниматься психологическими исследованиями — это получаемое от них удовольствие. И проводить их стоит, несмотря на случайные неудачи и разочарования, несмотря на огромное затраченное на них время. Это вновь и вновь подтверждается биографиями ведущих психологов-исследователей нашего времени. Примерами могут служить Элиот Аронсон — ученый из Стэнфорда — и Элизабет Лофтус из Вашингтонского университета.

Элиот Аронсон

Попросите психологов назвать имена самых выдающихся социальных психологов, и имя Элиота Аронсона будет в верхней части каждого списка. Среди психологов-исследователей он известен своей долгой и плодотворной работой в различных областях, начиная от фундаментальных исследований человеческих предрассудков, убеждений, когнитивного диссонанса до прикладных исследований (борьба со стереотипами, использование презервативов, экономия сил). Когда недавно *APA* вручила ему «Премию за крупный вклад в науку», последнее предложение документа о награждении гласило: «Видение Аронсоном социальной психологии как строгой науки, нужной человечеству, является маяком для будущих поколений ученых» (Award, 1999, p. 873). Он также знаком многим студентам — социальным психологам, так как, скорее всего, они занимаются по его учебнику. Работа Аронсона «Общественное животное» (Aronson, 1999a), получившая государственную награду в сфере средств коммуникации *APA,* выдержала восемь изданий и с каждым становится все совершеннее. Вы уже встречались с его именем в этой книге — именно он указал на различие между экспериментальным и жизненным реализмом (глава 3).

Аронсон был учеником выдающихся психологов. Будучи студентом университета, он случайно зашел на вводную лекцию, посвященную психологии предрассудков, которую читал известный психолог-гуманист Абрахам Маслоу. Тема привлекла внимание Аронсона — он вырос в рабочем городке Ривиэ в штате Массачусетс и нередко сталкивался с тем, что к нему, еврею, люди относятся, мягко говоря, предвзято. Поэтому Аронсон меняет специализацию на психолога и впоследствии близко сходится с Маслоу. Позже, получив степень магистра в Веслианском колледже, вместе с Дэвидом Мак-Клелландом (см. вставку 9.3) он уехал на запад — в Стэнфордский университет, куда он поступает в тот же год, когда Леон Фестингер начал в нем свою преподавательскую деятельность. Именно в это время Фестингер занимался разработкой своей известной теории когнитивного диссонанса, с которой вы вкратце

познакомились при обсуждении метода наблюдения участников исследования (вставка 12.1). Аронсон стал самым известным студентом Фестингера, а теория диссонанса стала центральной темой его дальнейших исследований. Аронсон и его ученики внесли большой вклад в разработку этой теории. Главным открытием Аронсона, сделанным в 1960 г., стала мысль о том, что когнитивный диссонанс наиболее силен и быстрее всего возникает, когда под угрозу ставится какой-либо важный аспект нашего представления о себе и когда мы делаем нечто (например, списываем на экзамене), противоречащее строгим убеждениям, которые мы считаем своими (например, «Я — честный человек») (Aronson, 1999b).

Как вы помните из главы 3, иногда идеи исследований приходят ученым-психологам во время будничных наблюдений, а иногда разрабатываются на основании дедуктивного анализа теорий. Аронсон находил свои идеи и тем и другим образом, нередко выводя гипотезу исследования в рамках теории диссонанса из наблюдения. Когда его спросили, как рождаются идеи его исследований, он указал на два источника: во-первых, «просто жизнь, в которой ты замечаешь любопытные вещи», и, во-вторых, «теории и пересмотренные теории» (Aronson, частная беседа, 2000). Однажды, например, он испытал чувство беспомощности, заметив, насколько неэффективны даже в эпоху СПИДа рациональные доводы при попытках убедить весьма неглупых студентов колледжа пользоваться презервативами. Расширенная теория диссонанса навела его на одну мысль. Предположив, что большинству людей не нравится чувствовать себя лицемерами, Аронсон разработал ряд исследований (например, Aronson, Fried, & Stone, 1991), в которых студенты участвовали в публичных дискуссиях о необходимости презервативов. Студенты, вынужденные аргументировать рациональную точку зрения, испытывали дискомфорт, поскольку сами презервативами не пользовались. Возникал диссонанс, который студенты устраняли, начиная пользоваться презервативами.

На протяжении всей долгой научной деятельности энтузиазм Аронсона не ослабевал. На вопрос о том, что помогает ему в исследованиях, он ответил просто и элегантно: «Три вещи. Во-первых, я страстно интересуюсь поведением людей в обществе, во-вторых, мне нравится делать открытия, способные улучшить жизнь людей, и наконец, сам процесс исследований — это большое удовольствие» (Aronson, частная беседа, 200).

Элизабет Лофтус

Работа Элизабет Лофтус воплотила совершенный синтез идеалов фундаментальных и прикладных исследований, которые описаны в главе 3. Как ученый она внесла огромный вклад в расширение наших знаний о фундаментальных процессах человеческой памяти и в то же время показала, как эти принципы соотносятся с явлениями обычной жизни, например воспоминаниями очевидцев происшествий. Сама Лофтус охарактеризовала этот синтез как наиболее интересный аспект своей исследовательской работы: «Я делаю интересные теоретические открытия и в то же время применяю их и другие психологические знания к практической жизни людей» (Loftus, частная беседа, 1993).

Элизабет Лофтус занимается изучением факторов, влияющих на долговременную память, и является, возможно, лучшим в мире специалистом по памяти очевидцев. Благодаря общепризнанной компетентности в этом вопросе она дала сви-

детельские показания в более чем 100 судебных процессах, для исхода которых показания очевидцев были решающими (Loftus, 1986). Как вы помните, одно ее исследование упоминалось в предыдущей главе, когда я предостерег вас от полного доверия к информации, почерпнутой из воспоминаний отдаленного прошлого. Лофтус вместе со своими коллегами и своей исследовательской группой провела десятки изысканий в области искажений воспоминаний действием информации, воспринятой между самим событием и его последующим припоминанием. Большинство из этих исследований было посвящено непосредственным воспоминаниям очевидцев. В последние годы внимание Лофтус привлек следующий эмпирический вопрос: могут ли воспоминания о сексуальных домогательствах, пережитых в детстве, сохраняться без изменений на протяжении многих лет. По ее мнению может возникнуть и возникает серьезное искажение таких воспоминаний.

Интерес Лофтус к исследованию памяти сформировался к концу курса ее обучения в Стэнфорде. Ее докторской диссертацией стал фундаментальный исследовательский проект, озаглавленный «Анализ структурных переменных, определяющих трудности в решении проблем, связанных с компьютерным телетайпом» Если у вас потемнеет в глазах, когда вы прочитаете такое название, Лофтус не обидится. Ближе к концу проекта она также «немного устала от всего этого» (Loftus & Ketchman, 1991, p. 5). К счастью, она изучала социальную психологию у Джона Фридмана, который пробудил в ней интерес к более практическим аспектам проблемы памяти. Лофтус написала об этом так:

> В последние шесть месяцев учебы в университете я проводила каждую свободную минуту в лаборатории Фридмана, подготавливая экспериментальные планы, работая с испытуемыми, составляя таблицы данных и анализируя результаты. Когда проект стал принимать определенные очертания и мы поняли, что близки к тому, чтобы сделать повое открытие о работе мозга, я начала считать себя нсихологом-исслсдователем. О! это были прекрасные слова — я могла разработать эксперимент, подготовить и провести его. Я впервые почувствовала, что стала ученым, и я абсолютно точно знала, что именно этим я и хотела заниматься в жизни,
>
> Loftus & Ketchman, 1991, p. 6

Есть одна черта, общая у Гибсон и Скпннера (см. главу 1) и у Аронсона, Лофтус и других психологов-исследователей. Павлов указал эту общую черту, когда посоветовал своим студентам *страатю* отдаваться работе. Если вы внимательно прочитаете выбранные мною цитаты в этой главе и главе 1, вы обнаружите слова именно об этом чувстве. Например, в высказываниях Гибсон и Скиннера встречается слово «замечательно». У Гибсон в первом пробном исследовании визуальной пропасти «все работало замечательно», когда Скиннер вернулся к застопорившемуся прибору для выдачи кусочков еды, он обнаружил «замечательную кривую». И Аронсон, и Лофтус, говоря о *«страстном* интересе» к социальному поведению людей и о *«прекрасных*словах» «психолог-исследователь», выразили то же самое сильное чувство.

В наших знаниях о законах поведения гигантские пробелы, а следовательно, перед людьми, движимыми страстью, подобной той, которая двигала Гибсон, Скиннером, Аронсоном, Лофтус и огромным количеством других исследователей, простирается широкое поле деятельности. Что может быть лучше, чем заниматься делом, постоянно приносящим радость и удовлетворение и помогающим понять сложность поведения и мыслительных процессов? Добро пожаловать!

ПРИЛОЖЕНИЕ *А*

Представление результатов психологических исследований

Прочитав это приложение, вы научитесь:

- составлять лабораторные отчеты в соответствии с нормами *APA]*
- готовить профессиональные презентации своих исследований:
  - в виде устного доклада;
  - в виде стенда.

Отчеты об исследованиях, подход *APA*

Вы вступаете в область, где господствуют навязчивые состояния. По мере изучения норм *APA* для составления лабораторных отчетов вам начнет казаться, что каждая пропущенная запятая, каждое слово, которое должно быть написано заглавными буквами, а написано строчными, каждое неподчеркнутое *Funw t* означает, что с вашим преподавателем что-то не в порядке. Еще будет казаться, что этот формат невозможно усвоить. Держитесь. Правила введены не просто так, а с определенной целью, а процедуру составления лабораторных отчетов вполне можно изучить. Кроме того, *APA* издает несложное «Руководство по публикациям» (РП) *{American Psychological Association,* 1994), в котором вы найдете ответы на все ваши вопросы по составлению отчетов. РП освободит вас от необходимости запоминать правила о запятых и подчеркиваниях.

Если вы еще не приобрели РП, но собираетесь изучать какие-либо дополнительные психологические предметы, вам следует сделать это как можно скорее. Данное приложение поможет вам узнать, как представлять результаты, по 368-страничное руководство — это более серьезное оружие в борьбе против ошибок.

Правила, определяющие единый формат представления результатов эмпирических исследований, нужны по двум причинам. Во-первых, по мере того как результаты исследования воспроизводятся в других лабораториях с другими участниками и, возможно, даже с несколько измененной процедурой, исследование накапливает внешнюю валидность. Чтобы в точности повторить исследование или внести определенные изменения в процедуру, читатель должен точно знать, каким образом было проведено оригинальное исследование. Унификация представления результатов исследования облегчает процесс его повторения. Во-вторых, строгий формат делает редакционный анализ статьи более эффективным. Каждый год

в десятки психологических журналов посылаются тысячи статей об исследованиях. Обычно, статьи просматриваются редактором, а затем посылаются одному или нескольким исследователям, компетентным в данном вопросе. Исследователи проводят экспертную оценку и посылают редактору критический обзор и мнение о том, стоит ли публиковать статью (большинство журналов отказывают в 70-80% случаев). Отсутствие заранее установленного формата серьезно осложнит работу редактора и экспертов. А вашему преподавателю методологии психологических исследований наличие строгого формата поможет правильно оценить ваш отчет.

Основные принципы

Я абсолютно уверен в том, что единственный способ научиться проводить психологические исследования — это провести несколько исследований. То же самое можно сказать об описаниях. Умение писать — это такой же навык, как и многие другие, который развивается при наличии практики и при условии, что прилагаемые усилия получают определенный отклик. Поэтому наиболее общая рекомендация, которую я могу дать, — это писать как можно чаще и просить остальных читать написанное вами. Если вы пишете лабораторный отчет, я могу посоветовать вам поискать человека, который прошел курс методологии исследований *и* получил высокую оценку по этому предмету. Попросите его прочитать ваш отчет и сделать критические замечания.

Стиль письма

Лабораторный отчет, составленный в стиле *АРА,* — это не американский роман, а поэтому такие литературные приемы, как сюжетная неопределенность, удерживающая интерес читателя, или опущенные подробности, вызывающие его любопытство, которые вполне на своем месте в занятиях по стилистике письма, совершенно неуместны в научной работе. При написании научных статей необходимо стремиться к полной ясности и простоте выражений. Конечно, легче сказать, чем сделать, но вопреки доле правды, которую содержит рис. *АЛ,* академическая работа не обязательно должна быть скучной и малопонятной.

Все хорошие отчеты и описания выполнены в соответствии с правилами грамматики. Грамматические ошибки создают двусмысленные и неуклюжие места в работе, поэтому первый шаг в обучении составлению лабораторных отчетов — это усвоение основных грамматических правил построения предложения. Существует множество справочников по английскому языку, кроме того, ряд принципов РП специально направлен на иллюстрацию типов предложений, используемых в научных отчетах. Одна особенно распространенная проблема научной манеры письма (кроме желания выглядеть объективно) — чрезмерное употребление пассивного залога. По возможности используйте активный залог — это способствует ясности и упрощает структуру предложений. Например:

Пассивный залог: Смитом было проведено исследование депривации сна.

Активный залог: Смит исследовал депривацию сна.

Дело не только в грамматике: очень часто начинающие составители лабораторных отчетов *АРА* неверно употребляют обычные слова. К примеру, термин «зна-

чимый» в обычном смысле подразумевает нечто важное, значительное, но если в отчете употребить этот термин в таком смысле, то читатель может неправильно вас понять, потому что в лабораторном отчете слово «значимый» означает «статистически значимый», а формальное определение этого термина связано с отклонением H_0. Читатель, встретивший это слово в лабораторном отчете, может решить, что его нужно понимать исходя из формального определения, даже если автор совсем не это имел в виду. Чтобы не создавать неопределенности, в лабораторных отчетах употребляйте этот термин только для обозначения статистической значимости.

Со словоупотреблением связаны еще две проблемы. Одна заключается в смешении слов «аффект» и «эффект», а вторая — в неправильном использовании множественного числа некоторых существительных, например существительного «data» (данные). Эти и другие проблемы рассматриваются в табл. *АЛ* и *А.2.*

Таблица А. 1

Правильное употребление терминов «аффект» и «эффект»

Одна из ошибок, чаще всего встречающихся в студенческих лабораторных отчетах, — это непонимание разницы между значениями слов «аффект» и «эффект».

Аффект означает эмоциональное состояние. Например:

- При высокой скорости показа испытуемые продемонстрировали увеличение аффекта.

Эффект означает результат некоторого события. Например:

- Нарастание скорости показа имело эффект снижения уровня запоминания









Рис. *АЛ.* Академическая работа?

(Кельвин и Гоббс © Watterson. Напечатано с разрешения Universal Press Syndicate. Все права защищены)

Таблица А.2

Множественное число слов латинского или греческого происхождения

Множественное число многих английских слов, встречающихся в научных работах, сохранило свою латинскую или греческую форму. Неспособность распознать множественное число может привести к грамматической ошибке — отсутствию согласования между существительным и глаголом. Наиболее распространенный пример — это английское слово «data» (данные), имеющее форму множественного числа, и слово «datum», имеющее форму единственного числа. Хотя термин «data» иногда, используется в речи в единственном числе, РП (р. 34) рекомендует употреблять его только во множественном числе.

Другие примеры:

Форма единственного числа	Форма множественного числа
analysis (анализ)	analyses (анализы)
criterion (критерий)	criteria (критерии)
hypothesis (гипотеза)	hypotheses (гипотезы)
phenomenon (феномен)	phenomena (феномены)
stimulus (стимул)	stimuli (стимулы)

Устранение искажений речи

Многие годы лингвисты и психологи, интересующиеся проблемами речи, исследуют взаимосвязь речи и мышления. Хотя природа этой взаимосвязи вызывает бурные споры, все согласны, что языковыми средствами можно усилить и закрепить определенные представления, в том числе стереотипные взгляды на поведение людей. Один из примеров, особо отмеченный в РП, — это язык, подразумевающий неравенство мужчин и женщин вследствие ориентации на определенный пол. В 1977 г. *APA* впервые разработала нормы для «несексистского», или «сексуально-нейтрального», языка. Авторы, отсылающие свои рукописи в журнал *APA,* должны употреблять сексуально-нейтральный язык.

Использование несексистского языка позволяет избежать двух проблем. Первая известна как «проблема обозначения». Она касается тех случаев, когда для обозначения и мужчин, и женщин используются слова, указывающие на мужской пол. Распространенный пример,— это использование слова «man» (человек, мужчина), когда подразумевается «person» (человек, человеческая единица), как, например, в высказывании: «Человек (man) с давних пор интересуется причинами психических заболеваний». Исследования показали, что даже если намерения авторов, употребляющих подобные выражения, абсолютно несексистские, читатели часто понимают их так, как будто они не имеет в виду женщин. Например, исследование Кидда (Kidd, 1971) показало, что когда испытуемые встречают слова «man» (человек, мужчина) или «his» (его) в высказываниях, не подразумевающих пола, они в 86% случаев воспринимают их как относящиеся только к мужчинам.

Ниже приведено несколько высказываний, иллюстрирующих проблему обозначения. Как вы увидите, эту проблему довольно легко решить переформулировав фразу, использовав множественное число или опустив слова, связанные с определенным полом. Примеры взяты из табл. 1 «Руководства по публикациям» (р. 54-60).

1. Ценность оказанной консультации лучше всего понимает сам клиент.

   Лучше сказать: Ценность оказанной консультации лучше всего понимают сами клиенты.

2. Стремление человека к знаниям.

   Лучше сказать: Стремление к знаниям.

3. Ученые-исследователи часто не в состоянии уделить достаточно времени своим женам и детям.

   Лучше сказать: Ученые-исследователи часто не в состоянии уделить достаточно времени своим супругам и детям.

Кроме проблемы обозначения при употреблении некоторых выражений, имплицитно содержащих значение пола, возникает «проблема оценки». Она появляется, когда слова, относящиеся к мужчинам и женщинам, подразумевают неравенство. Например, говоря о школьных спортсменах, автор может сказать «мужская баскетбольная команда» и «баскетбольная команда девушек» (лучше: «женская баскетбольная команда»). Выражения, подобные фразам «типично мужской» или «типично женский», или стереотипные определения одинакового для мужчин и женщин поведения (например, «амбициозный мужчина», но «агрессивная женщина» или «осторожный мужчина», но «робкая женщина») по этой же причине недопустимы.

Хотя в большинстве своем нормы для сексуально-нейтрального языка разработаны, чтобы не допустить ущемления прав женщин, один случай касается мужчин. Термин «материнская забота» нередко используется для описания опекающего отношения родителя к ребенку. Употреблять термины «материнская забота» по отношению к женщинам и «отцовская забота» для описания такого же поведения мужчин нельзя, так как выражение «отцовская забота» имеет другой смысл. *APA* рекомендует использовать сексуально-нейтральный термин «родительская забота», который позволяет понять, что мужчины также могут опекать ребенка.[1].

Искажение, связанное с обозначением пола, — это не единственная языковая проблема, указанная в РП. В руководстве также описаны языковые искажения, связанные с сексуальной ориентацией, расовой и этнической принадлежностью, возрастом и инвалидностью. Ниже приведены еще несколько примеров из табл. 1 «Руководства по публикациям».

1. Выборка состояла из 200 взрослых нетрадиционной сексуальной ориентации.

   Лучше сказать: Выборка состояла из 100 взрослых гомосексуалистов и 100 взрослых лесбиянок.

2. Мы изучали эскимосов...

   Лучше сказать: Мы изучали аборигенов Канады и Алеутских островов...

3. ...инвалид.

   Лучше сказать: ...человек с инвалидностью.

4. ...депрессивные.

   Лучше сказать: ...люди, страдающие депрессией.

[1] В английском языке противопоставление «материнской заботы» *(mothering)* **и «отцовской заботы»** *(fathering)* выражено гораздо сильнее, чем в русском. — *Примеч. перев.*

Академическая честность

В главе 2, в разделе, посвященном мошенничеству в науке, изложение было сосредоточено на проблеме фальсификации данных, но отмечался и вопрос плагиата (в конце главы было приведено соответствующее упражнение). В стандарте 6.22 этического кодекса 1992 г. специально указано на явление плагиата и говорится о том, что авторы «не выдают значительные части или отдельные детали чужих работ за свои собственные...» *(American Psychological Association,* 1992, p. 1609).

Основной вид плагиата — это непосредственное копирование информации из какого-либо источника без использования кавычек и ссылки на этот источник. Совет прост: ни при каких условиях не делайте этого, даже если информация получена из Интернета. Второй вид плагиата не столь очевиден, и студенты не всегда понимают, что занимаются плагиатом, когда должны были бы сослаться на источник. В этом случае студенты своими словами пересказывают чужие идеи. Данное в конце главы 2 упражнение посвящено именно этой проблеме. Каждый раз, когда вы используете идеи, термины, теории пли открытия, принадлежащие другим людям, необходимо ссылаться на их работы. Чтобы избежать плагиата, можно прямо цитировать работы других авторов и указывать, откуда взята цитата, или ссылаться на создателей излагаемых идей, опять-таки приводя источник. (Ниже приведены нормы *APA* по составлению списка литературы.)

Разное

Данные правила по составлению отчетов призваны облегчить процесс чтения создаваемых вами документов (ваш преподаватель может добавить или исключить некоторые пункты из этого списка):

- делайте двойные интервалы между строчками документа;
- делайте отступы (сверху, снизу и по краям) не менее 2,5 см;
- не сшивайте страницы, а используйте скрепки;
- не меняйте шрифт при наборе документа;
- не используйте **жирный шрифт** или *курсив* (слова, которые должны быть выделены курсивом в публикации, необходимо подчеркнуть в рукописи; см. пример документа ниже);
- сделайте себе копию рукописи и сдайте оригинал; не забудьте сохранить файл с вашей работой на диске.

Основные разделы лабораторного отчета

Лабораторный отчет в формате *APA,* содержащий описание эмпирического исследования, состоит из следующих разделов в указанном порядке:

- Титульный лист.
- Обзор.
- Введение.
- Метод.

- Результаты.
- Обсуждение.
- Список литературы.

За списком литературы обычно следуют страницы с авторскими примечаниями, сноски (если есть), таблицы, список рисунков и сами рисунки. Иногда в отчеты включают приложения (например, список стимулъного материала).

Далее рассказывается, как оформлять разделы лабораторного отчета. Обратите внимание на; а) табл. А3, в которой указаны наиболее распространенные (на мой взгляд) ошибки оформления каждого из разделов, и б) пример лабораторного отчета, приведенный в конце этого приложения.

Таблица А,3

20 основных ошибок оформления отчетов

На титульном листе:

1) отсутствие заголовка страницы и номера 1;
2) номер страницы не выровнен по правому отступу;
3) отсутствие выделения всего колонтитула заглавными буквами;
4) печать заглавными буквами всего названия работы;
5) заглавное написание «Ь» в «Running head» («Колонтитул»),

В обзоре:

1) написание слова «Обзор» заглавными буквами;
2) использование обычного отступа в первой строке;
3) использование более одного параграфа;
4) использование более 120 слов.

Во введении:

1) отсутствие повторения названия работы в начале раздела;
2) наличие слова «Введение» как названия раздела;
3) ошибки в оформлении ссылок.

В разделе «Метод»:

1) неправильное обозначение раздела (слово «метод» должно быть выровнено по центру);
2) отсутствие линии, подчеркивающей подзаголовки (например, «участники»);
3) печать раздела не сразу после введения, а с новой страницы.

В разделе «Результаты»:

1) неверное оформление заголовка раздела;
2) использование неправильных сокращений для статистических величин, таких как среднее арифметическое;
3) неверное представление заключительного статистического анализа.

В разделе «Обсуждение»:

1) неверное оформление заголовка раздела;
2) ошибки в оформлении ссылок.

Титульный лист

Как видно из приведенного ниже примера отчета, титульный лист оформляется строго определенным образом и состоит из следующих элементов.

Заголовок/номер страницы

Заголовок страницы помещается в верхнем правом углу листа и содержит полное название работы (два-три слова). Через пять пробелов после него идет цифра 1, означающая, что титульный лист является первой страницей рукописи. И заголовок страницы, и ее номер выровнены по правому краю (т. е. по краю правого отступа). Кроме страниц с рисунками (например, графиками, диаграммами, схемами), *все* страницы рукописи содержат заголовок и номер — это помогает определить, к какой работе относится страница, если по какой-либо причине она встретится отдельно от других. Для создания заголовка и номера страницы пользуйтесь стилем «Заголовок» шаблона.

Колонтитул

Откройте любую статью, напечатанную в журнале *APA,* и вы увидите, что в верхней части каждой страницы после первой находится «заголовок»: имя автора (четные страницы) или сокращенное название работы (нечетные страницы). Этот тип заголовка называется колонтитулом.

Чтобы поместиться в отведенной для него строке и передать смысл названия, колонтитул должен быть кратким, но достаточно информативным. Отметьте, что колонтитул и заголовок страницы — это разные вещи. Единственный случай, когда они могут совпадать, — это если первые несколько слов названия хорошо описывают содержание работы.

Колонтитул печатают в верхней части титульного листа, сразу под заголовком и номером страницы, и выравнивают по левому отступу. Его длина ограничена 50 знаками, включая буквы, пробелы между словами и знаки препинания. В случае примера с сеансом гипноза из главы 12 рукопись содержала колонтитул длиной 39 знаков:

Колонтитул: ПЕРЕЖИВАНИЯ; СОПУТСТВУЮЩИЕ ГИПНОЗУ НА СЦЕНЕ

Название/автор/место учебы или работы

Название статьи, имя человека, ее написавшего (или имена, если у статьи несколько авторов), и название колледжа, университета или другой организации, в которой работает автор, помещается в верхней половине страницы и выравнивается по центру. Первые буквы знаменательных слов названия должны быть заглавными. В примере с гипнотизмом (Crawford, Kitner-Triolo, Clarke, & Olesko, 1992) название работы было следующим:

Преходящие Позитивные и Негативные Переживания, Сопутствующие Гипнозу на Сцене[1]

[1] В России к оформлению работ предъявляются иные требования. — *Примеч. перев.*

Внимательно обдумайте название статьи (оно должно состоять из 10-12 слов). Название должно давать читателю ясное представление о содержании статьи и об изучаемых переменных. Например, названия экспериментальных исследований обычно имеют следующую структуру: «Влияние Хна У», что позволяет читателю определить зависимые и независимые переменные. Так, название исследования «Влияние ориентиров на местности и информации о местонахождении севера на определение направлений» говорит читателю о том, что участники будут выполнять задания на определение направлений (так будут находиться значения зависимой переменной) и что исследование будет включать две независимые управляемые переменные (наличие или отсутствие у участников информации о местонахождении севера и ориентиров на местности).

Второй распространенный вариант названия — повествовательное предложение, обобщающее результаты исследования. В пример можно привести следующее название: «Оказание помощи зависит от определенности ситуации и силы сочувствия очевидцев». Такое название сообщает читателю, что исследование включает управляемую независимую переменную (ясно ли из самой ситуации, что нужна помощь) и (возможно) субъектную переменную (способны ли очевидцы к сочувствию).

APA разработала особые правила для соавторства в исследовательском проекте. Как правило, нельзя стать соавтором без значительного вклада в проводимую работу. Участие в планировании исследования, помощь в его организации, сбор данных, работа по их анализу и высказывание своего мнения по поводу статьи, написанной другим исследователем, будут значительным вкладом в работу. Но ввода данных в таблицу и координации действий фонда субъектов будет недостаточно, чтобы войти в соавторы. В работах, имеющих несколько авторов, главным автором (т. е. тем, чье имя указано первым на титульном листе) должен быть исследователь, проделавший большую часть работы. Когда студенты ведут групповую работу над проектом, у них должен быть лидер, ответственность которого будет достаточно большой и который станет главным автором исследования.

Обзор

Этот раздел лабораторного отчета читают в первую очередь, а пишут в последнюю. Кроме того, большой объем данной в статье информации вынуждает многих читателей сначала просматривать обзоры, чтобы понять, стоит ли изучать статью более подробно, и нередко они ограничиваются *только* этим разделом. Обзоры полностью приводятся в базе данных *PsycINFO.*

Обзор содержит информацию четырех основных разделов, и поскольку он должен состоять из одного параграфа, не превышающего 960 знаков с пробелами (около 120 слов), его необходимо очень тщательно готовить и отбирать каждое слово. Для экономии места числа указываются в обзоре прописью, только если они начинают предложение. Первое предложение раздела наиболее информативно. В нем формулируют проблему и обычно знакомят читателей с исследуемыми индивидуумами и/или переменными. Ниже приведены несколько примеров первого предложения обзоров исследований, упомянутых в главе 12.

- Для описания сексуальных различий прикосновений людей было проведено наблюдение 4500 пар юношеского и более старшего возраста в общественный местах (Hall & Veccia, 1990, p. 1155).

- В ходе интервью с 22 участниками финансируемых университетами представлений была проведена оценка позитивных и негативных переживаний, сопутствующих публичным сеансам гипноза.
- Чтобы выяснить, может ли помещение больных в палату с видом на природу способствовать выздоровлению после холецистэктомии, были изучены истории болезней пациентов пригородной пенсильванской больницы с 1972 по 1981 г.

Кроме обобщающего содержание статьи первого предложения обзор включает информацию о методе, краткое изложение важнейших результатов (второе по важности предложение) и обычно заканчивается кратким выводом. Обзор располагается на с. 2 лабораторного отчета. Введение должно начинаться с новой страницы.

Введение

Во введении подробно описывается изучаемая проблема, проводится обзор литературы по данной теме и указываются известные и неизвестные аспекты проблемы. Неизвестные или неясные аспекты предполагают одно или несколько предсказаний, образующих гипотезу исследования. Таким образом, введение включает в себя, обычно в указанном порядке .формулировку проблемы, обзор соответствующей исследовательской литературы, а также одну или несколько гипотез, проверяемых в ходе исследования. В правильно написанном введении гипотезы естественным образом следуют из предшествующего описания проблемы и обзора литературы и приводятся в последнем абзаце. В слабом введении гипотезы без подготовки появляются в конце раздела и не имеют рационального основания. Проводимый во введении обзор литературы включает лишь те исследования, которые непосредственно связаны с работой, представленной в лабораторном отчете.

Введение начинается со с. 3 и представляет собой единственный раздел отчета, не имеющий заголовка, — оно начинается не со слова «введение», а с полного названия статьи.

Требования *APA для* оформления цитат

При обзоре исследований, связанных с изучаемой проблемой, вам потребуется приводить цитаты из этих работ. В РП представлены требования к оформлению цитат. Иногда имя автора является частью высказывания, описывающего конкретную работу, — в таком случае дата публикации указывается в скобках вслед за именем. Например:

> В исследовании Смита (1990) испытуемые оказывали помощь реже, если ситуация была менее очевидной.

Если имя автора не входит в такое высказывание, то в скобках указываются имя и дата публикации, разделенные запятой. Например:

> В одном из ранних исследований (Смит, 1990) испытуемые оказывали помощь реже, если ситуация была менее очевидной.

Если приводится прямая цитата, указывается также и номер страницы. Например:

Ситуация была неочевидной, когда «участники затруднялись определить, действительно ли человек, помогающим в проведении эксперимента, наступил па оголенный провод» (Смит, 1990, р. 23).

Или

В исследовании Смита (1990) ситуация была неочевидной, когда «участники затруднялись определить, действительно ли человек, помогающий в проведении эксперимента, наступил на оголенный провод» (р. 23).

Обратите особое внимание на следующий факт: каждая работа, упомянутая во введении (или в любой другой части отчета), должна быть представлена в полном списке литературы в конце статьи. Как оформлять такой список, вы узнаете из раздела «Список литературы». В «Руководстве по публикациям» приведены подробные указания к оформлению цитат, и оно ответит практически на любые ваши вопросы на данную тему. Кроме того, большинство ваших учебников по психологии, в том числе и этот, соответствуют требованиям *APA* — они могут послужить вам образцами.

Метод

Основной принцип написания раздела «Метод» заключается в том, что этот раздел должен быть достаточно подробным, чтобы другие исследователи могли прочитать его и воспроизвести исследование в своих лабораториях. Обратите внимание, что я написал «другие исследователи», а не «кто угодно» — вы можете считать, что читатель обладает специальными знаниями по методологии. Достаточно сказать, что участники были уравнены по переменным *X, Y* η Z; излагать процедуру уравнивания шаг за шагом нет необходимости. Аналогично можно указать вид использованного позиционного уравнивания (например, латинский квадрат 4x4) и не объяснять подробно, почему была необходима подобная процедура или как был построен квадрат.

В зависимости от исследования раздел «Метод» может состоять из нескольких подразделов, один из которых — «Участники» — является обязательным. Обычно он открывает данный раздел π представляет собой описание выборки людей или животных, принимавших участие в исследовании. Читатель этого подраздела должен иметь возможность определить общий тип участников исследования, их демографические характеристики, количество и способ их отбора.

За описанием участников обычно следует подраздел «Оборудование», если использовалось лабораторное оборудование, или «Материалы», если экспериментатор пользовался лишь бумажными материалами, такими как опросники или тесты личности. При описании стандартного оборудования обычно указывается производитель и номер модели. Приборы, сконструированные специально для исследования, должны быть описаны более подробно и, возможно, изображены на рисунке (как в приведенном ниже примере исследования). Защищенные авторскими правами опросники или анкеты личности можно просто перечислить с указанием источников.

Третий подраздел называется «Процедура» или иногда «Процедура и экспериментальный план». Он занимает большую часть раздела «Метод». В нем приводится описание и операциональное определение переменных, обсуждается план ис-

следования и особенности контроля, а также описывается, что именно происходило с участниками исследования.

В противоположность введению раздел, посвященный методу, не должен начинаться с новой страницы. Он идет сразу за введением н имеет заголовок «Метод», расположенный по центру (правила оформления подразделов см. в приведенном ниже примере исследования). Эти правила распространяются на разделы «Результаты» и «Обсуждение»

Результаты

В этом разделе отчета приводится краткое, но законченное описание результатов, а также данные статистики. Обычно абзац раздела о результатах включает в себя описание полученного результата и соответствующие данные описательного и заключительного статистического анализа. Не следует объяснять причины того, что прогноз оказался верным или наоборот, — это делается в разделе «Обсуждение».

Лучше всего представить результаты в том порядке, в котором следуют гипотезы во введении. Например, если введение заканчивается тремя гипотезами, раздел о результатах должен содержать абзацы, посвященные каждой из них, в том же порядке.

Описание данных: статистика

Как вы помните из главы 4, описательный статистический анализ обобщает данные, а заключительный показывает, можно ли отказаться от нулевой (нулевых) гипотезы (гипотез). Чаще всего приводятся следующие величины описательного статистического анализа: среднее арифметическое, стандартное отклонение и коэффициент корреляции. Статистика вывода включает проверку по критерию Стьюдента, коэффициент ^дисперсионного анализа и проверку по критерию хи-квадрат *(χ^1).* РП советует также приводить вычисление силы эффекта (р. 18). Процедуры нахождения этих величин изложены в приложении С, а сейчас вы должны иметь в виду две вещи. Во-первых, приняты определенные сокращения обозначений таких статистических показателей, как среднее значение по выборке (М) и стандартное отклонение *(SD).* Во-вторых, существует общее правило представления статистических данных, по которому в каждом конкретном абзаце величины описательной статистики предшествуют данным статистики вывода. Стандартный раздел «Результаты» обычно состоит из последовательности предложений следующего типа:

> Испытуемые, которым дали инструкции но построению визуальных образов, запомнили в среднем 16,5 слов *(SD* = 2,3), а те, кому было предложено просто повторять слова, запомнили в среднем 12,3 слов *(SD* = 3,1). Разница была значимой: *F(*\, 18)= 12,8;/;<0,01.

Краткий способ записи данных статистики вывода в целом такой же вне зависимости от типа анализа. Указываются тип анализа (например, *F),* степени свободы (1 для числителя коэффициента *Fu* 18 для знаменателя), найденное в результате анализа значение (12,8) и вероятность (например, /;<0,01). Обратите внимание, что, так же как *М* и *SD, Fu p* подчеркиваются. Это говорит о том, что при

публикации эти буквы будут выделяться курсивом. Это правило применяется и при оформлении раздела «Список литературы»: подчеркнутые слова при печати будут выделены курсивом (см. ниже).

Представление данных: таблицы и рисунки

Данные описательной статистики, особенно в случае факторных планов, нередко бывают настолько сложными, что их невозможно уместить в один абзац. Решением проблемы являются таблицы и/или рисунки. **Таблица** представляет собой форму организации данных по рядам и столбцам. Обычно в виде таблицы представляют значения среднего арифметического, стандартного отклонения или наборы значений корреляции. **Рисунком** может быть график, диаграмма, схема, чертеж прибора или фотография. Для каждого из них существуют особые правила оформления, которые можно изучить на примере приведенного ниже лабораторного отчета.

Иногда бывает нелегко решить, как представить данные — в виде таблицы или рисунка. Таблицы обычно используют, если важно указать точные значения среднего арифметического (на основании графика точные значения можно лишь предположить) или если данных так много, что график будет ими перегружен. А графики могут помочь более наглядно представить ситуацию. Более всего они подходят для отображения эффекта взаимодействия в факторных исследованиях. Следует учесть только, что представлять данные несколькими способами — и в виде таблицы, и в форме графика — не следует. Графики бывают различных видов; самые распространенные среди них — это линейные графики и гистограммы. Как показано в главе 7, линейные графики обычно (но не обязательно) строятся, если переменная, откладываемая по оси *X,* непрерывна. В случае дискретных переменных (например, пол), используют гистограммы.

Но недостаточно просто представить таблицу или график — в разделе «Результаты» необходимо сослаться на них и указать их важнейшие особенности. Для этого нужно обладать определенными навыками, поскольку подробно описать таблицу или график и указать все их аспекты нельзя. Так, график, изображающий кривую, полученную в эксперименте с памятью, может быть описан следующим образом:

> На рис. 1 показано, что запоминание было довольно сильным для первых слов списка, слабым для середины списка и практически полным для последних трех слов.

Как видно из приведенного ниже примерного отчета, таблицы и рисунки помещают не в основной текст, а в конец статьи.

Также обратите внимание на то, что в примере отчета заголовок таблицы располагается на той же странице, что и сама таблица, а названия рисунков — на отдельном листе. Это требование не случайно и разработано совсем не для того, чтобы можно было снизить вам оценку. Оно отражает процесс трансформации вашей работы в журнальную статью. Таблицы и названия рисунков набираются так же, как остальной текст, а сами рисунки представляют собой высококачественные фотографии и должны умещаться в определенные для них рамки (вот почему страницы с рисунками не содержат заголовков и номеров страницы). Поэтому рисунки и заголовки к ним располагаются отдельно друг от друга. Если ваша работа содержит несколько рисунков, не следует размещать каждый из заголовков на отдельном листе — их можно представить в виде списка на одной странице.

Обсуждение

Обсуждение представляет собой последний раздел, содержащий текстовые материалы, и служит для связи отчета в единое целое. Оно начинается с обобщения основных результатов, представляемых в соотношении с исходной гипотезой (гипотезами). Далее следует важнейшая часть раздела — интерпретация результатов. Проводимая оценка исследования включает установление взаимосвязи результатов с теоретическими вопросами, поднятыми во введении, и объяснение не оправдавшихся прогнозов. Также в ходе обсуждения рассматривается проблема альтернативных объяснений результатов. Как авторы статей, вы будете решать, какая из интерпретаций представляется наиболее рациональной, но иногда можно предположить несколько объяснений. Вам необходимо указать альтернативные варианты и объяснить, почему ваш — лучший из всех.

Завершает обсуждение важный элемент любой исследовательской программы: вопрос «что дальше?» Обычно исходя из результатов проведенного изыскания авторы могут предложить варианты новых экспериментов и будущих исследований. Это естественно следует из того факта, что исследование всегда отвечает на одни вопросы и поднимает другие.

Список литературы

В противоположность разделам «Метод», «Результаты» и «Обсуждение» список литературы начинается с новой страницы. Он содержит список работ, упоминаемых в лабораторном отчете, причем каждый элемент списка начинается с красной строки. В РП содержится полный набор правил оформления практически для всех видов работ. Ниже приведены наиболее распространенные примеры.

1. *Журнальная статья с одним автором:*

   Loftus, E. L. (1986). Ten years in life of an expert witness. Law and Human Behavior. 10.241-263.

2. *Журнальная статья с несколькими авторами:*

   Hall, J. A., & Veccia, E. M. (1990). More «touching» observations: New insights on men, women, and interpersonal touch. Journal of Personality and Social Psychology. 59,1155-1162.

3. *Статья в периодическом журнале:*

   Palmer, J. D. (1982, Ocober). Biorhythm bunkum. Natural History. 90-97.

4. *Книга*

   Skinner, B. F. (1953). Science and human behavior. New York: Free Press.

5. *Книга, выходящая не в первом издании:*

   Aronson, E. (1999). The social animal (8th ed.). New York: W. H. Freeman.

6. *Глава из изданной книги:*

   Weiss, J. M. (1977). Psychological and behavioral influences on gastrointestinal lesions in animal models. In J. D. Master & M. E. P. Seligman (Eds.), Psychopathology: Experimental models (pp. 232-269). San Francisco: Freeman.

7. *Электронные источники.*

В последние годы были разработаны и не так давно обновлены правила оформления ссылок на веб-сайты, электронные базы данных, электронную почту и т. д. Последние требования можно узнать на странице веб-сайта *APA,* посвященной данной теме: www.apa.org/jonrnals/webref.html.

Последнее замечание направлено на предотвращение распространенной ошибки: прежде чем закончить лабораторный отчет, необходимо сверить ссылки в тексте вашей работы со списком литературы. В особенности проверьте: а) все ли работы, упоминаемые в тексте лабораторного отчета, приведены в списке литературы и б) все ли работы, приведенные в списке литературы, упоминаются в тексте лабораторного отчета.

Устные доклады и стенды

Если вам повезет, то вы получите возможность представить результаты вашего исследования широкой публике. Презентация может быть адресована другим студентам, изучающим психологию, и быть предусмотрена обязательной программой обучения или проходить в более строгой форме на профессиональной конференции. В эпилоге говорилось о том, что доклады студентов включены, например, в расписание национальных заседаний *APA* и Американского психологического общества и региональных конференций (например, заседания «Среднезападной психологической ассоциации»). Кроме того, многие колледжи и университеты специально проводят конференции, чтобы дать студентам возможность представить свои работы.

Советы по проведению докладов

Презентации, проводимые на профессиональных психологических конференциях, бывают двух типов: устные или, чаще всего, в виде стендов. Устные презентации более традиционны: в порядке обычной сессии, длящейся 50 минут, несколько исследователей выступают с описаниями своих работ. Если вам придется выступать с устным докладом, воспользуйтесь следующими рекомендациями.

1. Вы будете ограничены жесткими временными рамками — придерживайтесь их. Обычно на выступление дается 12 минут, что дает возможность прослушать четыре доклада за сессию. Если вы займете 20-25 минут, последний по списку докладчик никогда вам этого не простит.
2. Подготовьте страничку тезисов, копии которой один из ваших друзей раздаст слушателям, когда вас будут представлять. Тезисы должны содержать заголовок работы, ваше имя и место учебы, гипотезу (гипотезы), сокращенное описание плана исследования, результаты (обычно в виде таблицы или графика) и краткие выводы. К сожалению, вместо того, чтобы внимательно слушать вас, большинство слушателей будут думать о сессии, которую они собираются посетить после этой, планировать обед или повторять собственный доклад, поэтому тезисы дают хорошую возможность присутствующим получить полезную информацию, которую они смогут просмотреть позже.

3. Если имеется аудио- или видеоаппаратура, воспользуйтесь ею. Например, если в зале, где проходит ваш доклад, есть диапроектор, подготовьте слайды, дублирующие таблицы и графики, представленные в тезисах. Продемонстрируйте их в нужные моменты вашего выступления и обратите внимание аудитории на наиболее яркие особенности. Зал может быть оснащен компьютерной техникой, поэтому научитесь пользоваться программой *PowerPoint* и с ее помощью постройте презентацию.
4. Подготовьте обычный лабораторный отчет, а затем сократите и переработайте его в устный доклад. Для этого выделите самые важные моменты и опустите мелкие детали. Например, если в лабораторном отчете указывается модель прибора, в устной презентации достаточно будет привести его название (например, «оперантная камера»). Слушателям должна быть ясна цель вашего исследования, важнейшие части плана и процедуры контроля, понятны результат])! и общие выводы. Доклад характеризуется более разговорной стилистикой по сравнению с более формальным лабораторным отчетом.
5. Отрепетируйте доклад перед друзьями, а лучше всего запишите на видеокассету, а затем попробуйте прочитать его еще несколько раз. Чтобы преодолеть неизбежную тревожность, подумайте о том, что по сравнению с работами других участников конференции ваша гораздо лучше.

Советы по оформлению стендов

Стендовые доклады позволяют представить гораздо больше исследовательских проектов в час, чем обычные устные, и, конечно, при этом докладчик не испытывает той тревожности, без которой не обходится на устных презентациях (возможно, это недостаток, ведь навык публичных выступлений — это очень важное качество). При этом также увеличивается вероятность встретить людей, чьи исследовательские интересы вам близки. В результате стендовой сессии вы примете участие в пяти-шести полезных разговорах, возможно, услышите новые, альтернативные интерпретации ваших достижений, разработаете несколько вариантов ответа на вопрос «что дальше?» и обменяетесь адресами электронной почты с другими исследователями.

Стендовое представление работ проходит в зале с выстроенными в ряд стендами. На одном из таких стендов вы повесите свой плакат, который покажет окружающим то, что вы обнаружили в ходе проведенного исследования. Также вам потребуется подготовить копии (обычно 50 штук) краткой версии вашей работы, чтобы раздавать желающим. Ниже приведены советы по оформлению стендов.

1. Очень важно правильно разместить информацию на стенде. Читатели должны всего за несколько минут понять, что произошло в ходе вашего исследования. Если их первая реакция — замешательство, они перейдут к другому стенду. На рис. *А.2* показано, как можно оформить стенд. На конференции, в которой вы примете участие, могут быть приняты определенные правила оформления стенда, так что обратите внимание на это. Например, на некоторых небольших конференциях все обзоры печатаются в особых буклетах, которые выдаются участникам. В таком случае не следует помещать обзор на стенд — начните с раздела «Введение» или «Формулировка проблемы».

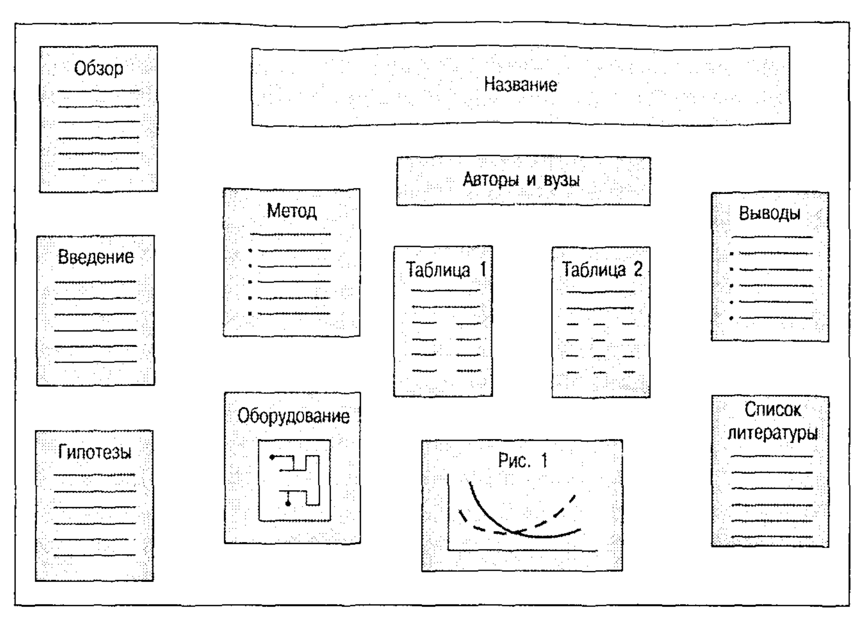


Рис. А.2. Примерный вид стенда

2. Материал на стенде должен быть логически организован. Читатели должны легко понимать цель вашего исследования, метод, результаты и общие выводы. Содержание плаката должно развиваться слева направо. Чтобы направлять внимание читателей, можно использовать стрелки.

3. Посетители должны иметь возможность читать и понимать информацию на стенде с расстояния двух метров. Высота букв названия должна быть около 4 см, а вашего имени и названия университета — немного меньше. Размер шрифта основных заголовков должен быть примерно 36 пунктов, заголовков рисунков — 24 пункта, а текста — не менее 18. Пользуйтесь шрифтами, которые легко читаются. Если вы используете на стенде минимальное количество цветов, это нормально. Слишком большое количество цветов может создать путаницу (поэтому не пытайтесь построить график, используя шесть различных цветов). Размещайте свою работу на подходящем фоне. Цвет стендов вполне подходит, но можно использовать и специальную бумагу. Возьмите с собой побольше кнопок, чтобы прикрепить страницы к стенду. На конференциях выдают кнопки, но их обычно на всех не хватает.

4. Краткость — сестра таланта, поэтому сведите объем к минимуму. Длинные абзацы замените на маркированные списки. Если кому-то потребуется более подробно узнать о некоторых аспектах исследования, вы сможете поговорить с ним.

5. Стенд — это наглядная форма представления работы, поэтому щедро используйте графики. Кроме таблиц и графиков на стенд можно поместить рисунки приборов, фотографии обстановки исследования и т. д. Но не забывайте, что каждый график должен отражать важную часть исследования.

ПРИЛОЖЕНИЕ В

Разработка опросов для исследований по психологии

В этом приложении вы научитесь:

- разрабатывать опросы для исследовательских проектов;
- создавать хороший инструментарий для оценки:
  - отношений;
  - знаний и памяти;
  - демографической информации.

Опрос — общее руководство по планированию

Как и любое другое исследование по психологии, опрос начинается с того, что на основании эмпирических вопросов разрабатывается гипотеза, которая далее проверяется путем сбора данных. Например, в колледжах или университетах подобные вопросы могут иметь конкретную направленность:

- Чувствуют ли себя студенты в безопасности, проходя по кампусу в ночное время?
- Едят ли студенты сбалансированную, питательную пищу?
- Являются ли преподаватели более либеральными, чем студенты?

Или они могут быть более общими:

- Довольны ли студенты своим обучением в университете?
- Если рассматривать поведение студентов в целом, можно ли охарактеризовать его как укрепляющее здоровье или подрывающее его?

Когда сформулированы эмпирические вопросы, необходимо тщательно спланировать опрос. Необходимо выбрать способ проведения опроса: в виде интервью с глазу на глаз, по телефону, с помощью письменных опросников или как комбинацию различных способов. О достоинствах и недостатках каждого из способов см. главу 12. Также необходимо выбрать процедуру построения выборки. Поскольку для проведения опроса важно иметь репрезентативную выборку, необходимо провести один из вариантов вероятностной выборки. Как вы помните из главы 4, они варьируются от простой случайной выборки до расслоенной и гнездовой. Если популяция небольшая (например, маленький кампус) и можно получить список

студентов, достаточно простой случайной или расслоенной выборки. Если необходимо точно отобразить в выборке некоторые важные подгруппы, имеющие различный вес в популяции, следует использовать процедуру расслоенной выборки. Например, если в кампусе 60% женщин и 40% мужчин, неплохо создать ту же пропорцию мужчин и женщин в выборке. Аналогично может возникнуть необходимость пропорционально отобразить количество первокурсников, второкурсников, студентов младших и старших курсов. Для больших популяций может потребоваться использовать одну из разновидностей кластерной выборки. Например, для длинного списка общеобразовательных предметов, которые предлагаются для изучения в течение одного семестра, можно сделать выборку и опросить всех студентов, посещающих отобранные предметы (с их личного согласия, конечно).

Поставив эмпирические вопросы и выбрав вид опроса и процедуру выборки, вы готовы приступить к размышлениям о том, какую информацию вам необходимо собрать в ходе этого опроса. Например, если вы хотите исследовать, какой образ жизни ведут студенты, здоровый пли нездоровый, вам необходимо составить список поведенческих актов людей и их отношения к различным вопросам — исследование будет направлено на изучение этого списка. Вы можете рассмотреть употребление алкоголя, питание, сексуальные отношения, курение, выполнение физических упражнений или несколько вопросов одновременно. Определив все эти параметры, можно приступить к разработке самого инструментария опроса.

Построение опросов

Предположим, что вы решили разработать и провести письменный опрос. Такие опросы обычно создаются с целью получения трех основных видов информации. Первый касается отношения опрашиваемых к каким-либо вопросам, второй включает в себя знания респондентов пли их способность запомнить определенные сведения, а третий представляет собой демографическую информацию — возраст, пол, социоэкоиомический статус 11 т. д.

Отношение

Отношение к определенным предметам пли вопросам — это общая тенденция думать, чувствовать и вести себя. Например, мы можем считать, что на работе к женщинам относятся несправедливо. Подобное отношение включает когнитивный компонент, в который могут входить определенные доводы в поддержку этого отношения (например различие в зарплате мужчин и женщин, выполняющих одинаковую работу), эмоциональный компонент (например, «Это неправильно и я недоволен этим»), который может различаться по силе, и склонность вести себя в соответствии с этими компонентами (например, «Я должен подписать петицию за уравнивание зарплаты»). Отношение часто оценивается с помощью опросов. Есть несколько общих правил для разработки таких опросов.

1. *Не изобретайте колесо* — по возможности используйте опубликованные опросы. Напечатаны сотни опросов, и если хороший опрос уже создан, вы

можете не затруднять себя созданием своего собственного, а получить разрешение на использование существующего. Опубликованные опросы можно найти с помощью базы данных *PsycINFO*. Например, если вас интересует отношение к исследованиям с использованием животных, поиск по данной теме без сомнений покажет исследование Плауса (Pious, 1996b), рассмотренное в качестве примера в главе 12. Далее вы можете написать Плаусу, спросить у него разрешение на использование опроса и попросить копию опросника. Другой неплохой источник — *Test Locator* в «Центре обмена информацией по оценке» *(ERIC, Clearinghouse on Assessment and Evaluation),* который находится по адресу http://ericae.net

2. *Используйте открытые вопросы с осторожностью.* Открытые вопросы могут быть полезны для получения самых разных ответов, в том числе таких, о которых исследователь не подозревает. Также они могут усилить у опрашиваемых чувство контроля за ситуацией при заполнении опросника. Но так как участники могут приводить разнообразные ответы, открытые вопросы трудно оценивать и выполнение такого исследования может потребовать заметно больше времени. Поэтому количество таких вопросов следует ограничить. Один неплохой способ — дать опрашиваемым возможность прокомментировать свои ответы на закрытые вопросы. Например, в конце списка закрытых вопросов в опроснике может быть отведено место для комментариев участников о любых вопросах, в особенности о тех, в случае которых были поставлены экстремальные оценки.

Еще один хороший способ применения открытых вопросов — в пробных исследованиях для определения вариантов ответов для будущего опросника, составленного из закрытых вопросов. Например, один их часто встречаемых открытых вопросов иногда называют «самой важной проблемой» (Schumann & Presser, 1996). В качестве примера можно привести следующий вопрос: «Как вы думаете, какие наиболее важные проблемы стоят сегодня перед этим университетом?»

Этот вопрос, заданный в пробном исследовании, вероятно, вызовет различные ответы, которые могут стать вариантами ответов, приведенными в заключительном варианте вопроса: «Что из перечисленного ниже является наиболее важной проблемой, стоящей сегодня перед этим университетом? (Также можно попросить оценить указанные проблемы по важности.)

_____Пожилые преподаватели.

_____Недостаточное компьютерное оснащение.

_____Авторитарный президент.

_____Плохие условия в общежитиях.

_____Недостаточная помощь в выборе профессии π т. д.».

Кроме того, некоторые вопросы могут быть «частично» открыты и включать перечень ответов, последним из которых идет графа «Прочее», в которую опрашиваемые могут вписывать свои ответы.

3. *Рассмотрите возможность использования шкалы Ликерта.* Как вы помните из главы 12, опросы, в которых используется шкала Ликерта, состоят из последовательности утверждений, а задание состоит в том, чтобы указать степень согласия или несогласия с этими утверждениями. Например, утверждение для опроса отношения к базовому учебному плану университета может быть сформулировано следующим образом:

   «Курс введения в психологию (PSY 100) должен быть обязательным для всех студентов».

   Варианты ответов из 5-баллыгой шкалы Ликерта могут быть следующими:

   _____полностью согласен;

   _____согласен;

   _____сомневаюсь (иногда больше подходит «нейтральное отношение»);

   _____не согласен;

   _____никак не согласен.

   Шкалу можно превратить в 7-балльную, добавив «абсолютно согласен» и «абсолютно не согласен». Явных преимуществ у 5- или 7-балльной (или более) шкалы не существует. 5-балльная шкала обычно обеспечивает достаточные различия между степенями согласия, но на практике она может стать de facto 3-балльной, принимая во внимание, что некоторые люди не склонны выбирать ответы, расположенные на обоих концах шкалы. 7-балльная шкала также дает 5 точек, если люди избегают точек экстремума, но введение дополнительных ответов может увеличить время, необходимое для завершения опроса. Общее правило состоит в том, чтобы не смешивать различные форматы. Если вы решили использовать 5-балльную шкалу Ликерта, везде используйте именно ее; не пользуйтесь 5- и 7-балльными шкалами в одном опросе.

   При использовании шкалы Ликерта формулируйте одни утверждения позитивно, а другие негативно. Например, в исследовании по оценке обучения последовательные утверждения могут быть сформулированы так:

   Преподаватель удерживал мое внимание во время лекций.

   Преподаватель составил такие вопросы для экзамена, которые адекватно проверяют знание изученного материала.

   Преподаватель был хорошо подготовлен к занятиям.

   Преподаватель оказывал дополнительную помощь вне лекций.

   Все эти ответы сформулированы позитивно, так что если студенту нравится преподаватель, он может пробежать опрос, не обращая внимания па отдельные вопросы, и со всем согласиться. Кроме того, ответы некоторых студентов подвержены искажению, которое социальные психологи называют «полное согласие в ответах». Оно состоит в тенденции соглашаться практически со всеми вопросами. Чтобы избежать таких проблем, опросы со шкалой Ликерта обычно содержат позитивные и негативные утверждения, что застав-

ляет участников внимательно читать и отвечать па каждое из них (Patten, 1998). Например, приведенные выше утверждения могут быть представлены следующим образом (два из них сформулированы позитивно, а два негативно):

Преподавателю редко удавалось удерживать мое внимание во время лекций.

Преподаватель составил такие вопросы для экзамена, которые адекватно проверяют знание изученного материала.

Преподаватель часто был неподготовлен к занятиям.

Преподаватель оказывал дополнительную помощь вне лекций.

4. *Внимательно подбирайте слова.* В главе 12 были кратко описаны несколько примеров возможных подводных камней. Необходимо избегать двусмысленностей, наводящих вопросов и искажений, а также не ставить два различных вопроса в виде одного. Приведем еще несколько советов (большинство примеров заимствовано из Fink, 1995).

   - Составляйте полные предложения.

     Плохо: Место жительства?

     Лучше: Как называется город, в котором **вы живете в настоящее время?**

   - Избегайте большинства аббревиатур.

     УЮК может быть как Университетом Южной Калифорнии, так и Университетом Южной Каролины.

   - Не используйте сленг и разговорные выражения, они выходят из моды; например, к моменту выхода этой книги выражение «информационный суперхайвей» уже устареет.

   - Не используйте жаргонные выражения, вы можете понимать разницу между «формирующей» и «суммарной» оценками (глава 10), но большинство участников опроса не понимают этого.

   - Не используйте негативную формулировку вопросов. Негативные утверждения труднее понимать, чем позитивные.

     Плохо: Считаете ли вы, что работники университета не должны иметь права осматривать вашу комнату в общежитии?

     Лучше: Должны ли работники университета иметь право осматривать вашу комнату в общежитии?

     — Используйте сбалансированные вопросы, а не такие, в которых отдается предпочтение той или иной точке зрения.

     Плохо: Поддерживаете ли вы использование животных при лабораторных занятиях студентов?

     Лучше: Поддерживаете ли вы или отвергаете использование животных при лабораторных занятиях студентов?

Знания и память

Общие правила, с которыми вы уже познакомились, распространяются также на опросы, призванные выявить знания опрашиваемых по определенной теме или их воспоминания о некотором событии. Но, к сожалению, при изучении базы знаний и/или памяти встречаются и особые проблемы.

1. *Не задавайте вопросов, на которые невозможно ответить из-за ограниченности памяти.* В вопросах о том, как часто участники совершают определенные поступки, задание слишком большого промежутка времени увеличивает шансы возникновения ошибок, связанных с памятью. Например, в опросе по изучению потребления алкоголя (предположим, что определение термина «выпивать» было разъяснено опрашиваемым):

   Плохо: Сколько раз вы выпивали за последний месяц?

   Лучше: Сколько раз вы выпивали за последнюю неделю?

   Конечно, промежуток времени будет зависеть от задаваемого вопроса. Если деятельность является относительно редкой (например, посещение зоопарка), небольшой интервал вызовет слишком много нулевых оценок. Один из способов избежать проблем с воспоминаниями — составлять списки. Например, в опросах, посвященных досугу, опрашиваемых можно попросить вспомнить последний месяц и указать, сколько раз они брали фильмы на прокат, ходили на прогулки, посещали филармонию и т. д.

2. *С осторожностью используйте ответ НЗ.* Очень часто честным ответом на вопрос о том, что знает человек, будет «Я не знаю». Поэтому вопросы, направленные на выявление знаний, часто содержат вариант ответа, называемый НЗ («Не знаю»). Некоторые специалисты не одобряют использование ответов НЗ, так как опрашиваемые могут злоупотреблять ими, консервативно отмечая НЗ, даже если они обладают некоторыми знаниями по данному вопросу. Результаты опросов с большим количеством ответов НЗ не слишком показательны. С другой стороны, отсутствие варианта НЗ может привести к тому, что опрашиваемые могут случайным образом выбрать ответ на вопрос, о котором они не имеют никакого представления (Schuman & Presser, 1996). Один из способов включить ответы НЗ в опрос и в то же время предотвратить их чрезмерное использование участниками — замаскировать вопросы, направленные на выявление знаний, предпослав им высказывания типа: «Постарайтесь предположить...» или «Не доводилось ли вам слышать или читать о том, что...» (примеры из Fink, 1995, p. 76). Кроме того, ответы НЗ должны использоваться только тогда, когда есть основания предполагать, что некоторые опрашиваемые не будут знать ответа на вопрос (Patten, 1998).

Демографическая информация

Использование вопросов, направленных на получение демографической информации, позволяет исследователю разделить результаты по категориям. Например, в опросе отношения студентов психологических специальностей к исследованиям с животными, приведенном в качестве примера в главе 12, Плаус (Pious, 1996b) просил участников указать свой пол, курс (например, второй) и ответить, собира-

ются ли они поступать в аспирантуру. Это позволило ему выяснить, что женщины чаще возражают против исследований с животными, чем мужчины, но также что отношение не изменяется систематически как функция от желания поступить в аспирантуру или курса.

Неплохо поместить демографические вопросы *в конец опроса.* Если вы с них начнете, участникам может стать скучно и они не будут отвечать на ключевые вопросы так хорошо, как должны. Также необходимо включать в опрос только те демографические категории, которые имеют значение для изучаемой гипотезы. Чем больше демографических вопросов вы зададите, тем длиннее будет опрос и выше риск, что участникам станет скучно. Кроме того, их могут раздражать подобные вопросы: запрос о получении некоторых демографических сведений (например, о доходе) может восприниматься как вмешательство в частную жизнь, даже если опрашиваемых заверили в конфиденциальности (Patten, 1998).

Есть определенные правила о том, как задавать демографические вопросы:

- Выясняя возраст, спрашивайте дату рождения.
- Выясняя годовой доход, приведите диапазоны зарплат, чтобы уменьшить опасения о вторжении в частную жизнь (например, 50 000 долларов или менее; 50 00 -100 000 долларов).
- Не допускайте пересечения вариантов ответов (например, как в вопросе о доходе, приведенном выше в качестве примера использования диапазонов).
- При проведении сравнения с другими опросами, задавайте те же демографические вопросы.
- Положитесь на профессионалов: например, вы можете позаимствовать вопросы из Управления переписи Соединенных Штатов (www.census.gov). Там, где новичок, к примеру, в графе «семейное положение» использует «женат/замужем, холост/не замужем, разведен(а) и вдовец/вдова», профессионал распознает неоднозначность категории «холост/не замужем» и укажет категории из переписи: «женат/замужем, живу отдельно, вдовец/вдова, разведена), никогда не был(а) женат/замужем» (Patten, 1998, стр. 27).

Советы

1. *Проводите пробное испытание опроса.* Если вы сами составляете вопросы, перепишите их по крайней мере дважды, а затем попросите двух друзей прочитать их, обращая внимание на ясность формулировок. Выберите таких друзей, которые не постесняются сказать вам, что конкретный вопрос просто ужасен. Когда вы поймете, что опрос достаточно хорош, проведите пробную проверку. Например, если вы составляете опрос в рамках курса по методам исследований, спросите у преподавателя, можно ли вам опробовать его на студентах, посещающих эти занятия. Затем попросите своих однокурсников: а) заполнить опросник и б) письменно покритиковать его. После чего выполните черновой анализ вопросов по полученным результатам. Допустим, к примеру, что опросник содержал следующий вопрос: «Сколько раз вы курили марихуану за последний месяц?»

   _____0-10;________11-20________21-30.

и все выбрали ответ «0-10». Это говорит о том, что категории слишком объемные. Исправленная версия должна содержать большее число ответов, в том числе категорию «0».

Внимательно составляйте введение. В опросах, которые рассылаются по почте, введение представляет собой сопроводительное письмо, в котором объясняются цели исследования, опрашиваемые заверяются в конфиденциальности и прилагаются все усилия, чтобы убедить участников заполнить опрос. Сам опрос может содержать вступительную инструкцию или несколько инструкций, предваряющих различные части опроса. Их также необходимо тщательно формулировать. Они должны быть кристально ясными и в то же самое время: а) не быть чрезмерно длинными и б) не ставить под сомнение интеллект опрашиваемого.

Тщательно прорабатывайте визуальное представление опроса. Если опрос будет выглядеть профессионально, то участники скорее поверят, что вы серьезно относитесь к исследованию, и в то, что это исследование заслуживает доверия. Например, между концом одного вопроса и началом другого должно быть оставлено достаточно места, опрос не должен содержать опечаток и быть качественно отпечатан (т. е. не пользуйтесь низкокачественными копировальными аппаратами).

Внимательно относитесь к порядку следования вопросов. Если отдельные вопросы некоторые участники могут счесть недопустимыми (например, о частоте употребления алкоголя или наркотиков, сексуальной жизни), поместите их в конец опроса. Если вы начнете с таких вопросов, опрашиваемые могут на них остановиться, а если они будут в конце, то участники потратят некоторое время, дойдя до них, и с большей охотой закончат опрос. Начните с таких вопросов, которые не являются личными, интересны и на которые легко ответить. Помещайте вопросы на одну тему в одной части опроса.

Вознаграждайте участников за ответы. Однажды я заполнил опрос об особенностях, которые я хотел бы найти в книгах по социальной психологии, потому, что получил за это подарочный сертификат книжного магазина Barnes & Noble на 5 долларов. Вероятно, вы не сможете позволить себе подобное вознаграждение, но можно рассмотреть другие способы увеличения процента ответов. Можно поступить просто и прикрепить к опросу карандаш или более сложно и вложить в него лотерейный билет для участия в розыгрыше относительно большого подарка. Плаус (Pious, 1996b), к примеру, проводя опрос студентов психологических специальностей об исследованиях с животными, подарил лазерный принтер психологическому факультету, а один из участвующих в опросе студентов по результатам лотереи выиграл 500 долларов (он также получил крупный грант от *National Science Foundation).*

Не забывайте об этике исследований. Следуйте требованиям исследовательского совета вашего учебного заведения, при необходимости подготовьте документ об обоснованном согласии (эту роль может играть сопроводительное письмо), предпримите все необходимые усилия, чтобы гарантировать конфиденциальность, а также не забудьте предоставить участникам возможность получить информацию о результатах опроса.

ПРИЛОЖЕНИЕ С

Использование статистических методов

Из данного приложении вы узнаете, как:

- определять степень взаимосвязанности двух переменных, вычисляя наиболее часто используемый коэффициент корреляции — пирсоново r;
- проводить простой регрессионный анализ
- двумя способами находить хи-квадрат (χ^2), т. е. проводить проверку номинальных данных;
- осуществлять проверку по критерию Стьюдента;

Разумное применение статистических методов

Один из важнейших навыков психолога-исследователя состоит в способности выбрать подходящий вид статистического анализа. В главах 4,7-9 вы уже частично познакомились с этим процессом. Из главы 4 вы узнали о различных видах шкалы измерений (номинальной, порядковой, интервальной и шкалы отношений) и о различиях между описательной статистикой и статистикой вывода. Также вы узнали, как вычислять некоторые общие величины, используемые в описательной статистике, например среднее арифметическое и стандартное отклонение, и познакомились с процедурой проверки гипотез. Главы 4,7 и 8 познакомили вас с проверкой по критерию Стьюдента и дисперсионным анализом *(ANOVA),* применяемым при заключительном анализе эксперимента. В главе 9 были описаны коэффициенты корреляции.

Из этого приложения вы узнаете, как выбрать подходящий вид анализа для полученных вами данных в соответствии с выбранным планом исследования, как проводить такой анализ и интерпретировать результаты с помощью статистических таблиц. Возможно, вы будете польюваться пакетами программ, например *Statistical Package for Social Sciences (SPSS),* которые предназначены для машинной статистической обработки результатов, но изучение примеров, приведенных в этом приложении, позволит вам лучше понять, как осуществляются подобные процедуры.

Какой вариант статистического анализа использовать — зависит от различных факторов, в том числе от: а) цели анализа: изучить взаимосвязь или провести срав-

нение, б) использованной шкалы измерений, в) особенностей плана исследования: например, является ли независимая переменная меж- или внутрисубъектной, и в некоторых случаях от г) размера выборки.

Если цель анализа — определить степень взаимосвязи двух изучаемых переменных, потребуется вычислить коэффициент корреляции. Самый общий из них — пирсоново г. Он используется, если измерения проводятся по интервальной шкале или шкале отношений. При использовании порядковой шкалы вычисляется «эр» Спирмена (обозначается г), показывающее степень взаимосвязи между двумя наборами порядковых данных. В случае номинальных данных также можно измерить взаимосвязь (коэффициент сопряженности — *С).* Ниже приведен пример вычисления пирсонова *г.* Как найти г и *С,* можно узнать из учебника по статистике.

Если целью анализа является сравнение двух или более экспериментальных условий, чтобы обнаружить существующие между ними *различия,* можно воспользоваться несколькими методами анализа. Приведенные в данном приложении примеры иллюстрируют некоторые наиболее общие подходы: проверку по критерию хи-квадрат (χ^2), [/-тест Мэнна—Уитни, проверку по критерию Стьюдента и дисперсионный анализ *(ANOVA).* Проверка по критерию χ-квадрат проводится, если получены номинальные данные, в случае порядковых данных используется [/-тест Мэнна—Уитни. Проверка по критерию Стьюдента и *ANOVA* требуют, чтобы данные были получены по интервальной шкале или шкале отношений.

Оценка взаимосвязей

Пример 1. Пирсоново r

Если обе переменные измеряются либо по интервальной шкале, либо по шкале отношений, их взаимосвязь можно оценить с помощью пирсонова *г.* Предположим, к примеру, что исследователь хочет определить взаимосвязь между количеством времени, бесполезно потраченного студентами, и их средним баллом. Средний балл варьируется от 0,0 до 4,0, а потраченное без пользы время — это количество часов, проводимых в неделю за определенными занятиями (например, просмотром мыльных опер). Для девяти студентов получены следующие данные.

>6 студента	**Бесполезно потраченное время Переменная *X***	X^2	Средний балл Переменная *Y*	Y^1	*XχY*
1	42	1764	1,8	3,24	75,6
2	23	529	3,0	9,00	69,0
3	31	961	2,2	4,84	68,2
4	35	1225	2,9	8,41	101,5
5	16	256	3,7	13,69	59,2
6	26	676	3,0	9,00	78,0
7	39	1521	2,4	5,76	93,6
8	19	361	3,4	11,56	64,6
Сумма	231	7293	22,4	65,50	609,7

Формула вычисления Пирсонова r:

$$r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}.$$

Шаг 1. *Вычислите все составляющие*

$$N\sum XY = (8)(609{,}7) = 4877{,}6;$$
$$\sum X\sum Y = (231)(22{,}4) = 5174{,}4;$$
$$N\sum X^2 = (8)(7293) = 58344;$$
$$(\sum X)^2 = (231)(231) = 53361;$$
$$N\sum Y^2 = (8)(65{,}5) = 524;$$
$$\sum(\sum Y)^2 = (22{,}4)(22{,}4) = 501{,}76.$$

Шаг 2. *Подставьте составляющие в формулу для r и вычислите его значение:*

$$r = \frac{4877{,}6 - 5174{,}4}{\sqrt{[58344 - 53361][524 - 501{,}76]}};$$

$$r = \frac{-296{,}8}{\sqrt{[4983][22{,}24]}};$$

$$r = \frac{-296{,}8}{\sqrt{[110821{,}9]}} = \frac{-296{,}8}{332{,}9} = -0{,}89.$$

Шаг 3. *Определите, является ли r значимым (т. е. отличается ли от нуля).* Узнать это можно с помощью табл. *D2* приложения *D,* в которой приведены «критические значения» (*cv - critical values)* величины r. Сначала определите степень свободы *(df-degree of freedom).* Для Пирсонова *rdf*= *N*- 2, где N—это количество пар оценок. В нашем примере *df*= 8 -2 =» 6. В строке таблицы, где *df=6,* вы найдете два критических значения: одно для уровня значимости, равного 0,05 *(cv* - 0,707), и второе для 0,01 *(cv* = 0,834). Если найденное значение *равняется или превышает* критическое значение, то можно отвергнуть нулевую гипотезу о том, что r = 0. Это означает возможность вывода о том, что корреляция является статистически значимой. В данном случае значение -0,89 является значимым при уровне значимости 0,01, так как оно больше критического значения, равного 0,834. Таким образом, вероятность того, что найденное значение корреляции (-0,89) является случайностью, очень мала (0,01 или 1 из 100). Является ли значение корреляции положительным или отрицательным, не важно — учитывается его абсолютное значение. Если вы внимательно изучите таблицу *D.2,* то сможете заметить один важный факт, касающийся корреляции. Если у вас всего несколько пар оценок (как в приведенном выше примере), корреляция должна быть довольно высокой, чтобы ее можно было охарактеризовать как значимую. Имея лишь несколько пар оценок, довольно легко случайно получить высокое значение корреляции. С другой стороны, при большом количестве пар оценок корреляция, кажущаяся весьма низкой, может тем не менее быть значимой.

Пример 2. Регрессионный анализ

Из главы 9 вы узнали о том, что наличие корреляции дает возможность делать прогнозы с помощью процедуры, называемой «регрессионным анализом». В ходе анализа определяется линия регрессии, которая служит основой для построения прогнозов. Ниже показан регрессионный анализ примера исследования взаимосвязи бесполезно потраченного времени и среднего балла. Формула для нахождения линии регрессии:

$$Y\text{-}a + bX,$$

где a = точка на оси Y;

Ъ » наклон кривой;

X = известное значение;

Y = предполагаемое значение.

Шаг 1. *Вычислите все составляющие:*

$$b = r\frac{s_y}{s_x},$$

где *г* = пирсоново *г* - -0,89;

s = стандартное отклонение (см. табл. 4.4);

s_y = 0,63;

s_x = 9,43;

$$b = -0{,}89\frac{0{,}63}{9{,}43} = (0{,}89)(0{,}07) = -0{,}06.$$

$$a = \bar{Y} - b\bar{X},$$

где $\bar{Y}$ = среднее арифметическое для Y = 2,80;

$\bar{X}$ = среднее арифметическое для X = 28,88;

a = 2,80 - (-0,06)(28,88) = 4,53.

Шаг 2. *Подставьте значения для точки на оси Y и наклона кривой в формулу регрессии:*

$$Y = a + bX = 4{,}53 + (-0{,}06)X = 4{,}53 - 0{,}06X.$$

Шаг 3. *Используйте формулу для предположений.*

Если Патрик потратил без пользы 40 часов, каков его предположительный средний балл?

$$Y = 4{,}53 - 0{,}06.$$

$$X = 4{,}53 - (0{,}06)(40) = 4{,}53 - 2{,}40 = 2{,}13.$$

Если Патриция потратила 20 часов без пользы, каков ее предположительный средний балл?

7=4,53-0,06.

X= 4,53 - (0,06)(20) = 4,53-1,20 = 3,33.

Оценка различий

Пример 3. Степень согласия χ^2

Если отчет о данных включает в себя указание числа попаданий некоторого явления в определенную категорию, то становится ясно, что использована номинальная шкала измерений. Чтобы определить, является ли частота этого события систематической величиной или просто случайностью, требуется провести заключительную проверку номинальных данных. Проверка по критерию хи-квадрат (χ^2) — это наиболее общая статистическая процедура, используемая в случае номинальных данных. Есть две разновидности этой процедуры, различающиеся в зависимости от того, сколько наборов значений получено, один или два. Процедуру, используемую в случае одного набора значений, иногда называют «степень согласия χ^2», так как с ее помощью определяют, отклоняется ли частота, полученная в исследовании, от частоты, ожидаемой по теории вероятности или в соответствии с конкретной предсказываемой моделью.

В качестве примера предположим, что у студентов складывается мнение, будто используемые преподавателем задания на выбор ответа содержат искажения: пять вариантов ответа не одинаково часто являются правильными. Им кажется, что варианты *б, в* и *г* являются правильными чаще, чем *а* или *д*. Преподаватель заинтересовался этим вопросом и решил оценить все задания на выбор ответа, использованные в предыдущем семестре. Было подсчитано, сколько раз выбирается каждый из вариантов ответа. Если искажение отсутствует, это значение должно быть примерно одинаковым для всех вариантов. Таким образом, нулевая гипотеза говорит, что для каждого из вариантов ожидаемая частота выбора будет одинаковой. Общее количество вариантов в заданиях на выбор ответа равняется 400, а следовательно, ожидаемая частота (Е) для каждого из вариантов равняется 400/5 = 80. В действительности получены следующие значения частоты:

Вариант *а:*	62.
Вариант *б:*	85.
Вариант *в:*	78.
Вариант *г:*	111.
Вариант *д:*	64.
Итого:	400.

Формула для вычисления степень согласия χ^2:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}.$$

Шаг 1 *Вычислите все составляющие*

О-Е:

62 - 80 = -18;
85 - 80 = +5;
78 - 80 = -2;
111 — 80 = +31;
64 - 80 = -16.

(О-Е)²:

$(-18)^2 = 324;$
$(+5)^2 = 25;$
$(-2)^2 = 4;$
$(+31)^2 = 961;$
$(-16)^2 = 256.$

Шаг 2 ***Подставьте** составляющие в формулу для* χ^2 *и найдите его значение:*

$$\chi^2 = \sum\frac{(O-E)^2}{E} =$$

$$= \frac{324}{80} + \frac{25}{80} + \frac{4}{80} + \frac{961}{80} + \frac{256}{80} =$$

$$= 4{,}05 + 0{,}31 + 0{,}05 + 12{,}01 + 3{,}20 =$$

$$= 19{,}62$$

Шаг 3 *Определите, значимо ли найденное значение* χ^2.

В табл. D.3 приведены «критические значения» для χ^2. Степень свободы для χ^2 в случае одной выборки равняется количеству категорий минус 1, в данном случае 4 (5 - 1 = 4). В строке таблицы для *df*= 4 находятся два критических значения: одно для уровня значимости 0,05 *(cv* = 9,49), а второе — для уровня 0,01 (о> = 13,28). Найденное значение превышает оба критических значения, а следовательно, коэффициент χ^2 значим при уровне значимости 0,01. Преподаватель сделает вывод, что в распределении вариантов правильных ответов для задания на выбор присутствует некоторое искажение. Варианты *а* и Э действительно используются реже других.

Пример 4. χ^2 — два набора значений

В исследованиях по психологии чаще всего находят χ^2, если получено более одного набора значений частоты. Наиболее распространенный вариант — использование двух различных групп испытуемых и разделение членов каждой группы на одну или несколько категорий в зависимости от изучаемого вопроса. Предположим, к примеру, что исследователь хочет определить, влияют ли половые различия на выбор определенных предметов в качестве профилирующих. Чтобы выяснить количество мужчин и женщин, выбравших психологию, биологию и математику, изучаются заявления, подаваемые поступающими студентами (т. е. проводится архивная процедура). В следующей таблице, называемой «факторной таблицей», приведены полученные результаты:

	Психология	Биология	Математика	Сумма в ряду
Мужчины	13	17	20	50
Женщины	24	16	10	50
Сумма в столбце	37	33	30	100

Нулевая гипотеза говорит, что на выбор предметов не влияют половые различия. Для проверки нулевой гипотезы для полученных значений частоты находится χ^2. Формула для χ^2 в случае двух наборов значений:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}.$$

Шаг 1. *Вычислите значения ожидаемой частоты (Е):*

E = [(сумма в ряду)(сумма в столбце)]/общая сумма.

Для ячейки «мужчины, выбравшие психологию» (м, п):

$E_{мп}$ - (50 x 37)/100 - 1850/100 = 18,5.

Для остальных ячеек:

$E_{мб}$ - (50 x 33)/100 - 1650/100 = 16,5;

$E_{ИМ}$ - (50 x 30)/100 - 1500/100 = 15,0;

$E_{жсп}$ - (50 x 37)/100 - 1850/100 - 18,5;

$E_{жб}$ - (50 x 33)/100 = 1650/100 = 16,5;

$E_{жм}$ - (50 x 30)/100 = 1500/100 - 15,0.

Шаг 2 *Вычислите $(O - E)^2$ для каждой ячейки:*

$(O-E)^2$: м,п $\rightarrow (13 - 18{,}5)^2 = 30{,}25$;

м,б $\rightarrow (17 - 16{,}5)^2 = 0{,}25$;

м,м $\rightarrow (20 - 15{,}0)^2 = 25{,}00$;

ж,п $\rightarrow (24 - 18{,}5)^2 = 30{,}25$;

ж,б $\rightarrow (16 - 16{,}5)^2 = 0{,}25$;

ж,м $\rightarrow (10 - 15{,}0)^2 = 25{,}00$.

Шаг 3 *Подставьте составляющие в формулу для χ^2 и найдите его значение:*

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} =$$

$$= \frac{30{,}25}{18{,}5} + \frac{0{,}25}{16{,}5} + \frac{25{,}00}{15{,}0} + \frac{30{,}25}{18{,}5} + \frac{0{,}25}{16{,}5} = \frac{25{,}00}{15{,}0} =$$

$= 1{,}64 + 0{,}02 + 1{,}67 + 1{,}64 + 0{,}02 + 1{,}67 =$

$= 6{,}66$

Шаг 4 *Определите, значимо ли найденное значение* χ^2.

Степень свободы для χ^2 в случае двух наборов значений равняется: (кол-во рядов - 1)(кол-во столбцов - 1) = (3 -1)(2 - 1) = 2.

В табл. *D.3* в строке для *df=2* критические значения равны 5,99 (уровень значимости 0,05) и 9,21 (уровень 0,01). Найденное значение, равное 6,66, больше первого критического значения, но меньше второго, а следовательно, коэффициент χ^2 значим для уровня значимости 0,05, но не для 0,01. Половые различия влияют на выбор профилирующего предмета: значительно больше женщин показали интерес к психологии и значительно больше мужчин — к математике. И женщины, и мужчины показали одинаковый интерес к биологии.

Пример 5. Проверка по критерию Стьюдента — независимые группы

В исследованиях с одной независимой переменной, принимающей всего два значения, различия между двумя наборами оценок часто оцениваются с помощью критерия Стьюдента. Как вы помните из главы 7, есть два основных вида проверки по критерию Стьюдента — они различаются в зависимости от того, являются ли группы испытуемых, на которых получены оценки, независимыми или нет. Независимые группы образуются, если испытуемые распределяются случайным образом или используется субъектная переменная, например пол или возраст. Такие экспериментальные планы требуют проверки по критерию Стьюдента для независимых групп. Проверка по критерию Стьюдента для зависимых групп (иногда называемых «коррелированными» группами) используется, когда одни и те же испытуемые изучаются при обоих условиях эксперимента или если разные группы участников определенным образом взаимосвязаны — либо посредством процедуры уравнивания (см. пример 7), либо с помощью естественного уравнивания, которое наблюдается в случае сравнения родителей с детьми. Ниже представлен простой способ вычисления коэффициента Стьюдента для независимых групп. Для вычислений используется дисперсия, которая, как вы помните из главы 4, является важным показателем изменчивости набора оценок, и находится ее квадратный корень, что дает стандартное отклонение. Более подробно о вычислении стандартного отклонения см. табл. 4.4. В целом, при проверке по критерию Стьюдента для независимых групп сравниваются различия между группами с дисперсией в пределах каждой группы. Исследователи надеются, что различия между группами будут огромными, тогда как изменчивость в пределах групп будет небольшой.

Предположим, исследователь проводит простой эксперимент с памятью и с помощью случайного распределения сформировал две группы испытуемых. Одна группа изучает список из 25 слов при скорости показа 2 с на слово, а другая — при скорости 4 с на слово. Ниже приведено количество слов, запомненных пятью членами каждой группы.

N субъекта	2 с/слово (X_1)	N субъекта	2 с/слово (X_2)
1	14	6	18
2	11	7	23

3	12	8	19
4	17	9	17
5	13	10	22
Сумма	67		99
η	5		5
Среднее арифметическое	13,4		19,8
Стандартное отклонение	2,3		2,6
Дисперсия	5,3		6,7

В ходе проверки по критерию Стьюдента разница между двумя средними арифметическими, полученными по результатам эксперимента, делится на «стандартную ошибку различия» — предположительную оценку того, как сильно должны расходиться значения среднего арифметического при влиянии случайных факторов или возникновении ошибки. Исследователь надеется на то, что числитель будет большим, а знаменатель — маленьким, а следовательно, будет большим значение t. В таком случае различия между средними арифметическими будут больше, чем ожидается при воздействии только случайных факторов.

Формула для вычисления коэффициента Стьюдента для независимых групп следующая:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}.$$

Шаг 1. *Найдите все составляющие:*

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}};$$

$$t = \frac{13,4-19,8}{\sqrt{\left[\frac{(5-1)5,3+(5-1)6,7}{5+5-2}\right]\left[\frac{1}{5}+\frac{1}{5}\right]}};$$

$$t = \frac{-6,4}{\sqrt{\left[\frac{21,2+26,8}{8}\right][0,2=0,2]}};$$

$$t = \frac{-6,4}{\sqrt{[(6)(0,4)]}} = \frac{6,4}{1,55};$$

$$t = -4,13.$$

Шаг 2. *Подставьте составляющие в формулу и вычислите значение t.*

Шаг 3. *Определите, является ли найденное значение t значимым.*

Степень свободы для коэффициента Стьюдента для независимых групп равняется:

$$(и_1 + я_2 - 2) = (5 + 5 - 2) = 8.$$

В табл. *D5* представлен список критических значений для оценки результатов проверки по критерию Стьюдента. В строке, где *df*=>8, критические значения равняются 2,31 (уровень значимости 0,05) и 3,36 (уровень значимости 0,01). Найденное значение 4,13 превосходит оба из них (знак минус не учитывается), а следовательно, *t* значимо для уровня 0,01. В данном случае будет разумно отвергнуть нулевую гипотезу и заключить, что у испытуемых, которым демонстрировали слова с разной скоростью, запоминание различается.

Шаг 4. *Оцените силу эффекта.*

Как вы помните из главы 4, обычно исследователи не только выясняют, являются ли различия между значениями среднего арифметического статистически значимыми, но также определяют относительную силу эффекта, вызываемого экспериментальным воздействием. При проверке по критерию Стьюдента сила эффекта равняется величине изменчивости зависимой переменной, вызываемой независимой переменной (Cohen, 1988). Существуют различные способы оценки силы эффекта; один из наиболее распространенных — коэново *d*. Чтобы его вычислить, необходимо найти разность между значениями среднего арифметического и разделить ее на предполагаемое стандартное отклонение в популяции, значение которого находится для обоих групп:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s}.$$

Чтобы найти предполагаемое стандартное отклонение в популяции, необходимо сложить значения дисперсии для двух групп и из полученного значения извлечь квадратный корень. Получаем:

$$s = \sqrt{5{,}3 = 6{,}7} = \sqrt{12} = 3{,}46.$$

Тогда сила эффекта равняется:

$$d = \frac{13{,}4 - 19{,}8}{3{,}46} = -1{,}85.$$

Что означает такой результат? Согласно общим принципам, предложенным Коэном (процитировано в Spatz, 1997), силу эффекта можно разделить на малую (около 0,2), среднюю (около 0,5) и большую (около 0,8). По этому стандарту 1,85 — это очень большой эффект (знак минус можно не учитывать, он лишь показывает, какое из значений среднего арифметического стоит первым в числителе). Таким образом, увеличение ско-

рости показа с 2 до 4 секунд на слово в данном примере имеет заметное влияние на запоминание.

Примечание, анализ силы эффекта можно провести для второго вида проверки по критерию Стьюдента, с которым вы сейчас познакомитесь, а также для различных видов *ANOVA* (анализа дисперсии). Чтобы изучить конкретные процедуры, обратитесь к учебнику по статистике.

Пример 6. Проверка по критерию Стьюдента — зависимые группы

Как было отмечено ранее, проверка по критерию Стьюдента для зависимых групп проводится, если используются планы с повторяемыми измерениями и уравненными группами, а независимая переменная принимает два значения. Каждая пара оценок будет иметь определенные внутренние взаимосвязи, поскольку получена от: а) испытуемых, имеющих некоторое сходство между собой, или б) одних и тех же испытуемых. Так же как коэффициент Стьюдента для независимых групп, в случае зависимых групп этот коэффициент отражает отношение действительных различий между значениями среднего арифметического к изменчивости в пределах каждого условия. Процедура включает вычисление корреляции между двумя наборами оценок и подстановку этого значения в формулу для нахождения коэффициента Стьюдента. В приведенном ниже примере используется упрощенная формула, позволяющая проводить непосредственное вычисление коэффициента Стьюдента без предварительного нахождения пирсонова *г.*

Предположим, что исследователь для сравнения двух способов обучения компьютерной грамотности — курса для самостоятельного изучения и обычного лекционного курса — использует план с уравненными группами. 10 студентов в каждой группе были уравнены по среднему баллу и коэффициенту вербального интеллекта. Так, пара испытуемых N1 (см. ниже) состоит из двух человек, имеющих примерно одинаковые средний балл и уровень интеллекта. Зависимая переменная может принимать максимальное значение, равное 35. Ниже приведены данные исследования и результаты предварительного анализа, включающего вычисление D И D^2 ДЛЯ каждой пары оценок.

Пары испытуемых	Курс для самостоятельного изучения	Лекции	D	D^2
1	26	18	8	64
2	31	22	9	81
3	26	21	5	25
4	28	20	8	64
5	22	17	5	25
6	22	15	7	49
7	23	21	2	4
8	29	20	9	81
9	24	19	5	25
10	24	23	1	1
Сумма N- 10	255	196	59	419
Среднее арифметическое	25,5	19,6		

Формула для вычисления коэффициента Стьюдента для зависимых групп:

$$t = \frac{\sum D}{\sqrt{\left[\frac{N\sum D^2 - (\sum D)^2}{N-1}\right]}}.$$

Шаг 1. *Найдите все составляющие:*

$$\sum D = 8 + 9 + ... + 1 = 59;$$

$$(\sum D)^2 = (59)^2 = 3481;$$

$$\sum D^2 = 64 + 81 + ... + 1 = 419.$$

Шаг 2. *Подставьте составляющие в формулу и вычислите значение t.*

$$t = \frac{\sum D}{\sqrt{\left[\frac{N\sum D^2 - (\sum D)^2}{N-1}\right]}};$$

$$t = \frac{59}{\sqrt{\left[\frac{(10)(419) - 3481}{9}\right]}};$$

$$t = \frac{59}{\sqrt{\left[\frac{709}{9}\right]}} = \frac{59}{\sqrt{78,8}} = \frac{59}{8,9};$$

$$t = 6,63.$$

Шаг 3. *Определите, является ли найденное значение t значимым.*

Степень свободы для коэффициента Стьюдента для независимых групп равняется количеству пар оценок минус 1, в данном случае df= 10 - 1 = 9. Снова воспользуйтесь табл. *D.5.* Для строки, в которой df = 9, критические значения равны 2,26 (уровень значимости 0,05) и 3,25 (уровень значимости 0,01). Найденное значение (6,63) превышает оба критических значения, а следовательно, *t* значим при уровне значимости 0,01. Таким образом, курс для самостоятельного изучения эффективнее традиционного лекционного курса.

ПРИЛОЖЕНИЕ D

Статистические таблицы

В данном приложении вы найдете таблицы, используемые, когда необходимо принять статистически обоснованное решение:

Таблица *Di.* Случайные числа.

Таблица *D2.* Критические значения для пирсонова г.

Таблица *D3.* Критические значения для хи-квадрат (χ^2).

Таблица *DA.* Критические значения для íZ-теста Мэнна—Уитни.

Таблица *D5.* Критические значения для распределения *t* (двусторонний тест).

Таблица *D6.* Критические значения для распределения *F.*

Таблица D. 1

Случайные числа

03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 26 16 80 45	60 11 14 10 95
97 74 24 67 62	42 81 14 57 20	42 53 32 37 32	27 07 36 07 51	24 51 79 89 73
16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59	88 97 54 14 10
12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21	88 26 49 81 76
55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	0618 44 32 53	23 83 01 30 30
16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	87 35 20 96 43	84 26 34 91 64
84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	21 76 33 50 25	83 92 12 06 76
63 01 63 78 59	16 95 55 67 19	98 10 50 71 75	12 86 73 58 07	44 39 52 38 79
33 21 12 34 29	78 64 56 07 82	52 42 07 44 38	15 51 00 13 42	99 06 02 79 54
57 60 86 32 44	09 47 27 96 54	49 17 46 09 62	90 52 84 77 27	08 02 73 43 28
18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	06 76 50 03 10	55 23 64 05 05
26 62 38 97 75	84 16 07 44 99	83 11 46 32 24	20 14 85 88 45	10 93 72 88 71
23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	32 98 94 07 72	93 85 79 10 75
52 36 28 19 95	50 92 26 11 97	00 56 76 31 38	80 22 02 53 53	86 60 42 04 53
37 85 94 35 12	83 39 50 08 30	42 34 97 96 88	54 42 06 87 98	35 85 29 48 39
70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	17 76 37 13 04	07 74 21 19 30
56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	70 33 24 03 54	97 77 46 44 80
99 49 57 22 77	88 42 95 45 72	16 64 36 16 00	04 43 18 66 79	94 77 24 21 90
16 08 15 04 72	33 27 14 34 09	45 59 34 68 49	12 72 07 34 45	99 27 72 95 14
31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	52 85 66 60 44	38 68 88 11 80

Окончание табл. D. 1

68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52	68 07 97 06 57
74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76	15 54 55 95 52
27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	96 46 92 42 45	97 60 49 04 91
00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 65 04 26	11 04 96 67 24
29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	34 25 20 57 27	40 48 73 51 92
16 90 82 66 59	83 62 64 11 12	67 19 00 71 74	60 47 21 29 68	02 02 37 03 31
11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	76 70 90 30 86	38 45 94 30 38
35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	16 92 53 56 16	02 75 50 95 98
38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	40 01 74 91 62	48 51 84 08 32
31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	00 52 43 48 85	27 55 26 89 62
66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	76 83 20 37 90	57 16 00 11 66
14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
20 46 78 73 90	97 51 40 14 02	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
64 19 58 97 79	15 06 15 93 20	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
05 26 93 70 60	22 35 85 15 13	11 05 65 09 68	76 83 20 73 90	57 16 00 11 66
07 97 10 88 23	09 98 42 99 64	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
68 71 86 85 85	54 87 66 47 54	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
26 99 61 65 53	58 37 78 80 70	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
14 65 52 68 75	87 59 36 22 41	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
05 26 93 70 60	22 35 85 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
07 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 0577 09 51
68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43
05 26 93 70 60	22 35 85 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
07 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 05 77 09 51
68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43
17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	44 95 27 36 99	02 96 74 30 83
90 26 59 21 19	23 52 23 33 12	96 93 02 18 39	07 02 18 36 07	25 99 32 70 23
41 23 52 55 99	31 04 49 69 96	10 47 48 45 88	13 41 43 89 20	97 17 14 49 17
60 20 50 81 69	31 99 73 68 68	35 81 33 03 76	24 30 12 48 60	18 99 10 72 34
91 25 38 05 90	94 58 28 41 36	45 37 59 03 09	90 35 57 29 12	82 62 54 65 60
34 50 57 74 37	98 80 33 00 91	09 77 93 19 82	74 94 80 04 04	45 07 31 66 49
85 22 04 39 43	73 81 53 94 79	33 62 46 86 28	08 31 54 46 31	53 94 13 38 47
09 79 13 77 48	73 82 97 22 21	05 03 27 24 83	72 89 44 05 60	35 80 39 94 88
88 75 80 18 14	22 95 75 42 49	39 32 82 22 49	02 48 07 70 37	16 04 61 67 87
90 96 23 70 00	39 00 03 06 90	55 85 78 38 36	94 37 30 69 32	90 89 00 76 33

Таблица D.2

Критические значения для Пирсонова r

Значение альфа (а)			Значение альфа (а)		
df	0,05	0,01	*df*	0,05	0,01
1	0,997	1,00	16	0,468	0,590
2	0,950	0,990	17	0,456	0,575
3	0,878	0,959	18	0,444	0,561
4	0,811	0,917	19	0,433	0,549
5	0,755	0,875	20	0,423	0,537
6	0,707	0,834	25	0,381	0,487
7	0,666	0,798	30	0,349	0,449
8	0,634	0,767	35	0,325	0,418
9	0,602	0,735	40	0,304	0,393
10	0,576	0,708	45	0,288	0,372
11	0,553	0,694	50	0,273	0,354
12	0,532	0,661	60	0,250	0,325
13	0,514	0,641	70	0,232	0,302
14	0,497	0,623	80	0,217	0,283
15	0,482	0,606	90	0,205	0,267
			100	0,195	0,254

Источник: Fisher, R.A., & Yates, F. (1963) Statistical tables for biological, agricultural, and medical research (6th ed.). Table VII. Edinburg: Oliver & Boyd.

Таблица D.3

Критические значения для хи-квадрат (χ^2)

Значение альфа (а)			Значение альфа (а)		
df	**0,05**	**0,01**	*df*	**0,05**	**0,01**
1	3,84	6,64	16	26,30	32,00
2	5,99	9,21	17	27,59	33,41
3	7,82	11,34	18	28,87	34,80
4	9,49	13,28	19	30,14	36,19
5	11,07	15,09	20	31,41	37,57
6	12,59	16,81	21	32,67	38,93
7	14,07	18,48	22	33,92	40,29
8	15,51	20,09	23	35,17	41,64
9	16.92	21,67	24	36,42	42,98
10	18,31	23,21	25	37,65	44,31
11	19,68	24,72	26	38,88	45,64
12	21,03	26,22	27	40,11	46,96
13	22,36	27,69	28	41,34	48,28
14	23,68	29,14	29	42,56	49,59
15	25,00	30,58	30	43,77	50,89

Источник: Fisher, R.A., & Yates, F. (1963) Statistical tables for biological, agricultural, and medical research (6th ed.). Table IV. Edinburg: Oliver & Boyd.

Таблица D.4

Критические значения для U-теста Мэнна—Уитни[1]

Односторонний тест **α= 0,01 (обычный шрифт)** **α= 0,005 (жирный шрифт)**	**Двусторонний тест** **α= 0,02 (обычный шрифт)** **α= 0,01 (жирный шрифт)**

$N_2\backslash N_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–	–	–	–	0	0	0	0	0	0	1	1
													–	**–**	**–**	**–**	**–**	**–**	**0**	**0**
3	–	–	–	–	–	–	0	0	1	1	1	2	2	2	3	3	4	4	4	5
							–	**–**	**0**	**0**	**0**	**1**	**1**	**1**	**2**	**2**	**2**	**2**	**3**	**3**
4	–	–	–	–	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
					–	**0**	**0**	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**5**	**5**	**6**	**6**	**7**	**8**
5	–	–	–	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
				–	**0**	**1**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
6	–	–	–	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
				0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**9**	**10**	**11**	**12**	**13**	**15**	**16**	**17**	**18**
7	–	–	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
			–	**0**	**1**	**3**	**4**	**6**	**7**	**9**	**10**	**12**	**13**	**15**	**16**	**18**	**19**	**21**	**22**	**24**
8	–	–	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
			–	**1**	**2**	**4**	**6**	**7**	**9**	**11**	**13**	**15**	**17**	**18**	**20**	**22**	**24**	**26**	**29**	**30**
9	–	–	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
			0	**1**	**3**	**5**	**7**	**9**	**11**	**13**	**16**	**18**	**20**	**22**	**24**	**27**	**29**	**31**	**33**	**36**
10	–	–	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
			0	**2**	**4**	**6**	**9**	**11**	**13**	**16**	**18**	**21**	**24**	**26**	**29**	**31**	**34**	**37**	**39**	**42**
11	–	–	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
			0	**2**	**5**	**7**	**10**	**13**	**16**	**18**	**21**	**24**	**27**	**30**	**33**	**36**	**39**	**42**	**45**	**48**
12	–	–	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
			1	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27**	**31**	**34**	**37**	**41**	**44**	**47**	**51**	**54**
13	–	0	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
		–	**1**	**3**	**7**	**10**	**13**	**17**	**20**	**24**	**27**	**31**	**34**	**38**	**42**	**45**	**49**	**53**	**56**	**60**
14		0	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
		–	**1**	**4**	**7**	**11**	**15**	**18**	**22**	**26**	**30**	**34**	**38**	**42**	**46**	**50**	**54**	**58**	**63**	**67**
15	–	0	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
		–	**2**	**5**	**8**	**12**	**16**	**20**	**24**	**29**	**33**	**37**	**42**	**46**	**51**	**55**	**60**	**64**	**69**	**73**
16	–	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
		–	**2**	**5**	**9**	**13**	**18**	**22**	**27**	**31**	**36**	**41**	**45**	**50**	**55**	**60**	**65**	**70**	**74**	**79**
17	–	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
		–	**2**	**6**	**10**	**15**	**19**	**24**	**29**	**34**	**39**	**44**	**49**	**54**	**60**	**65**	**70**	**75**	**81**	**86**
18	–	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
		–	**2**	**6**	**11**	**16**	**21**	**26**	**31**	**37**	**42**	**47**	**53**	**58**	**64**	**70**	**75**	**81**	**87**	**92**
19	–	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
		0	**3**	**7**	**12**	**17**	**22**	**28**	**33**	**39**	**45**	**51**	**56**	**63**	**69**	**74**	**81**	**87**	**93**	**99**
20	–	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114
		0	**3**	**8**	**13**	**18**	**24**	**30**	**36**	**42**	**48**	**54**	**60**	**67**	**73**	**79**	**86**	**92**	**99**	**105**

[1] Чтобы *U*, вычисленное на основании полученных данных, **было значимым,** его значение должно равняться или быть ниже, чем значение, приведенное в таблице. Прочерки в ячейках таблицы означают, что для установленного уровня значимости решение принять невозможно. *Источник:* Kirk, R. E (1984). Elementary Statistics. Belmont, CA: Brooks/Cole.

Окончание табл. D.4

Односторонний тест α= 0,01 (обычный шрифт) α= 0,005 (жирный шрифт)											**Двусторонний тест α= 0,02 (обычный шрифт) α= 0,01 (жирный шрифт)**									
$N_2 \backslash N_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0
2	–	–	–	–	0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
					–	**–**	**–**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**2**	**2**	**2**	**2**
3	–	–	0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
	–	**–**	**–**	**–**	**0**	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**	**5**	**5**	**6**	**6**	**7**	**7**	**8**
4	–	–	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	–	**–**	**–**	**0**	**1**	**2**	**3**	**4**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**11**	**12**	**13**	**13**
5	–	0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	–	**–**	**0**	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**	**17**	**18**	**19**	**20**
6	–	0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	–	**–**	**1**	**2**	**3**	**5**	**6**	**8**	**10**	**11**	**13**	**14**	**16**	**17**	**19**	**21**	**22**	**24**	**25**	**27**
7	–	0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
	–	**–**	**1**	**3**	**5**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**	**34**
8	–	1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
	–	**0**	**2**	**4**	**6**	**8**	**10**	**13**	**15**	**17**	**19**	**22**	**24**	**26**	**29**	**31**	**34**	**36**	**38**	**41**
9	–	1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	–	**0**	**2**	**4**	**7**	**10**	**12**	**15**	**17**	**20**	**23**	**26**	**28**	**31**	**34**	**37**	**39**	**42**	**45**	**48**
10	–	1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
	–	**0**	**3**	**5**	**8**	**11**	**14**	**17**	**20**	**23**	**26**	**29**	**33**	**36**	**39**	**42**	**45**	**48**	**52**	**55**
11	–	1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	–	**0**	**3**	**6**	**9**	**13**	**16**	**19**	**23**	**26**	**30**	**33**	**37**	**40**	**44**	**47**	**51**	**55**	**58**	**62**
12	–	2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	–	**1**	**4**	**7**	**11**	**14**	**18**	**22**	**26**	**29**	**33**	**37**	**41**	**45**	**49**	**53**	**57**	**61**	**65**	**69**
13	–	2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	–	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**33**	**37**	**41**	**45**	**50**	**54**	**59**	**63**	**67**	**72**	**76**
14	–	2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
	–	**1**	**5**	**9**	**13**	**17**	**22**	**26**	**31**	**36**	**40**	**45**	**50**	**55**	**59**	**64**	**67**	**74**	**78**	**83**
15	–	3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	–	**1**	**5**	**10**	**14**	**19**	**24**	**29**	**34**	**39**	**44**	**49**	**54**	**59**	**64**	**70**	**75**	**80**	**85**	**90**
16	–	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	–	**1**	**6**	**11**	**15**	**21**	**26**	**31**	**37**	**42**	**47**	**53**	**59**	**64**	**70**	**75**	**81**	**86**	**92**	**98**
17	–	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	–	**2**	**6**	**11**	**17**	**22**	**28**	**34**	**39**	**45**	**51**	**57**	**63**	**67**	**75**	**81**	**87**	**93**	**99**	**105**
18	–	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
	–	**2**	**7**	**12**	**18**	**24**	**30**	**36**	**42**	**48**	**55**	**61**	**67**	**74**	**80**	**86**	**93**	**99**	**106**	**112**
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	–	**2**	**7**	**13**	**19**	**25**	**32**	**38**	**45**	**52**	**58**	**65**	**72**	**78**	**85**	**92**	**99**	**106**	**113**	**119**
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	–	**2**	**8**	**13**	**20**	**27**	**34**	**41**	**48**	**55**	**62**	**69**	**76**	**83**	**90**	**98**	**105**	**112**	**119**	**127**

Таблица D.5

Критические значения для распределения *t* (двусторонний тест)

Значение df	альфа (α)		Значение альфа (α)		
	0,05	0,01	*df*	0,05	0,01
1	12,71	63,66	18	2,10	2,88
2	4,30	9,93	19	2,09	2,86
3	3,18	5,84	20	2,09	2,85
4	2,78	4,60	21	2,08	2,83
5	2,57	4,03	22	2,07	2,82
6	2,45	3,71	23	2,07	2,81
7	2,37	3,50	24	2,06	2,80
8	2,31	3,36	25	2,06	2,79
9	2,26	3,25	26	2,06	2,78
10	2,23	3,17	27	2,05	2,77
11	2,20	3,11	28	2,05	2,76
12	2,18	3,06	29	2,05	2,76
13	2,16	3,01	30	2,04	2,75
14	2,15	2,98	40	2,02	2,70
15	2,13	2,95	60	2,00	2,66
16	2,12	2,92	120	1,98	2,62
17	2,11	2,90	?	1,96	2,58

Источник: Fisher, R.A., & Yates, F. (1963) Statistical tables for biological, agricultural, and medical research (6[dl] ed.). Table III. Edinburg: Oliver & Boyd.

Таблица D. 6

Критические значения для распределения *F*

Степени свободы (знаменатель)	Степени свободы (числитель)								
№	1	2	3	4	5	6	7	8	9
1	161,4	199,5	215,7	224,6	230,0	234,0	236,8	238,9	240,5
	4052,40	**4999,5**	**5403**	**5625**	**5764**	**5859**	**5928**	**5981**	**6022**
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
	98,50	**99,00**	**99,17**	**99,25**	**99,30**	**99,33**	**99,36**	**99,3**	**99,39**
3	10,31	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
	34,12	**30,82**	**29,46**	**28,71**	**28,24**	**27,91**	**27,67**	**27,49**	**27,35**
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,06
	21,20	**18,00**	**16,69**	**15,98**	**15,52**	**15,21**	**14,98**	**14,80**	**16,66**
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
	16,26	**13,27**	**12,06**	**11,39**	**10,97**	**10,67**	**10,46**	**10,29**	**10,16**
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
	13,75	**10,92**	**9,78**	**9,15**	**8,75**	**8,47**	**8,26**	**8,10**	**7,98**
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
	12,25	**9,55**	**8,45**	**7,85**	**7,46**	**7,19**	**6,99**	**6,84**	**6,72**
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
	11,26	**8,65**	**7,59**	**7,01**	**6,63**	**6,37**	**6,18**	**6,03**	**5,91**
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
	10,56	**8,02**	**6,99**	**6,42**	**6,06**	**5,80**	**5,61**	**5,47**	**5,35**
10	4,96	4,10	3,29	3,48	3,33	3,22	3,14	3,07	3,02
	10,04	**7,56**	**6,55**	**5,99**	**5,46**	**5,39**	**5,20**	**5,06**	**4,94**
11	4,84	3,98	3,24	3,36	3,20	3,09	3,01	2,95	2,90
	9,65	**7,21**	**6,22**	**5,67**	**5,32**	**5,7**	**4,89**	**4,74**	**4,63**
12	4,75	3,89	3,20	3,26	3,11	3,00	2,91	2,85	2,80
	9,33	**6,93**	**5,95**	**5,41**	**5,06**	**4,82**	**4,64**	**4,50**	**4,39**
13	4,67	3,81	3,16	3,18	3,03	2,92	2,83	2,77	2,71
	9,07	**6,70**	**5,74**	**5,21**	**4,86**	**4,62**	**4,44**	**4,30**	**4,19**
14	4,60	3,74	3,13	3,11	2,96	2,85	2,76	2,70	2,65
	8,86	**6,51**	**5,56**	**5,04**	**4,69**	**4,46**	**4,28**	**4,14**	**4,03**
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
	8,68	**6,36**	**5,42**	**4,89**	**4,56**	**4,32**	**4,14**	**4,00**	**3,89**
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
	8,53	**6,23**	**5,29**	**4,77**	**4,44**	**4,20**	**4,03**	**3,89**	**3,78**
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49
	8,40	**6,11**	**5,18**	**4,67**	**4,34**	**4,10**	**3,93**	**3,79**	**3,68**
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
	8,29	**6,01**	**5,09**	**4,58**	**4,25**	**4,01**	**3,84**	**3,71**	**3,60**

Продолжение табл. D.6

Степени свободы (числитель)									
10	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	*x*
241,9	243,9	245,9	248,9	249,9	250,1	251,1	252,2	254,3	254,6
6056,90	6106	6157	6209	6235	6261	6287	6313	6339	6366
19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
99,40	**99,42**	**99,43**	**99,45**	**99,46**	**99,47**	**99,47**	**99,48**	**99,49**	**99,50**
8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
27,23	**27,05**	**26,87**	**26,69**	**26,60**	**26,50**	**26,41**	**26,32**	**26,22**	**26,13**
5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
14,55	**14,37**	**14,20**	**14,02**	**13,93**	**13,84**	**13,75**	**13,65**	**13,56**	**13,46**
4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
10,05	**9,89**	**9,72**	**9,55**	**9,47**	**9,50**	**9,29**	**9,20**	**9,11**	**9,02**
4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7,87	**7,72**	**7,56**	**7,40**	**7,31**	**7,23**	**7,14**	**7,06**	**6,97**	**6,88**
3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
6,62	**6,47**	**6,31**	**6,16**	**6,07**	**5,99**	**5,91**	**5,82**	**5,74**	**5,65**
3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
5,81	**5,67**	**5,52**	**5,36**	**5,28**	**5,20**	**5,12**	**5,03**	**4,95**	**4,86**
3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
5,26	**5,11**	**4,96**	**4,81**	**4,73**	**4,65**	**4,57**	**4,48**	**4,40**	**4,31**
2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
4,85	**4,71**	**4,56**	**4,81**	**4,33**	**4,25**	**4,17**	**4,08**	**4,00**	**3,91**
2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
4,54	**4,40**	**4,25**	**4,10**	**4,02**	**3,94**	**3,85**	**3,78**	**3,69**	**3,60**
2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
4,30	**4,16**	**4,01**	**3,86**	**3,78**	**3,70**	**3,62**	**3,54**	**3,45**	**3,36**
2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
4,10	**3,96**	**3,82**	**3,66**	**3,59**	**3,51**	**3,43**	**3,34**	**3,25**	**3,17**
2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
3,94	**3,80**	**3,66**	**3,51**	**3,43**	**3,35**	**3,27**	**3,18**	**3,09**	**3,00**
2,54	2,54	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
3,80	**3,67**	**3,52**	**3,37**	**3,29**	**3,21**	**3,13**	**3,05**	**2,96**	**2,87**
2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
3,69	**3,55**	**3,41**	**3,26**	**3,18**	**3,10**	**3,02**	**2,93**	**2,84**	**2,75**
2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
3,59	**3,46**	**3,31**	**3,16**	**3,08**	**3,00**	**2,92**	**2,83**	**2,75**	**2,65**
2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
3,51	**3,37**	**3,23**	**3,08**	**3,00**	**2,92**	**2,84**	**2,75**	**2,66**	**2,57**

Продолжение табл. D.6

Степени свободы (знаменатель)	Степени свободы (числитель)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42
	8,18	**5,93**	**5,01**	**4,50**	**4,17**	**3,94**	**3,77**	**3,63**	**3,52**
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
	8,10	**5,85**	**4,94**	**4,43**	**4,10**	**3,87**	**3,70**	**3,56**	**3,46**
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37
	8,02	**5,78**	**4,87**	**4,37**	**4,04**	**3,81**	**3,64**	**3,51**	**3,40**
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34
	7,95	**5,72**	**4,82**	**4,31**	**3,99**	**3,76**	**3,59**	**3,45**	**3,35**
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32
	7,88	**5,66**	**2,76**	**2,26**	**3,94**	**3,71**	**3,54**	**3,41**	**3,30**
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30
	7,82	**5,61**	**4,72**	**4,22**	**3,90**	**3,67**	**3,50**	**3,36**	**3,26**
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28
	7,77	**5,57**	**4,99**	**4,18**	**3,85**	**3,63**	**3,46**	**3,32**	**3,22**
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
	7,72	**5,53**	**4,98**	**4,14**	**3,82**	**3,59**	**3,42**	**3,29**	**3,18**
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25
	7,68	**5,49**	**4,60**	**4,11**	**3,78**	**3,56**	**3,39**	**3,26**	**3,15**
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24
	7,64	**5,45**	**4,57**	**4,04**	**3,75**	**3,53**	**3,36**	**3,23**	**3,12**
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22
	7,60	**5,42**	**4,54**	**4,04**	**3,73**	**3,50**	**3,33**	**3,20**	**3,09**
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
	7,56	**5,39**	**4,51**	**4,02**	**3,70**	**3,47**	**3,30**	**3,17**	**3,07**
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
	7,31	**5,18**	**4,31**	**3,83**	**3,51**	**3,29**	**3,12**	**2,99**	**2,89**
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
	7,08	**4,98**	**4,13**	**3,65**	**3,34**	**3,12**	**2,95**	**2,82**	**2,72**
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96
	6,85	**4,79**	**3,95**	**3,48**	**3,17**	**2,96**	**2,79**	**2,66**	**2,56**
∞	3,84	3,00	2,60	3,37	2,21	2,10	2,01	1,94	1,88
	6,63	**4,61**	**3,78**	**3,32**	**3,02**	**2,80**	**2,64**	**2,51**	**2,41**

Окончание табл. D.6

Степени свободы (числитель)									
10	12	15	20	24	30	40	60	120	*x*
2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
3,43	**3,30**	**3,15**	**3,00**	**2,92**	**2,84**	**2,76**	**2,67**	**2,58**	**2,49**
2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
3,37	**3,23**	**3,09**	**2,94**	**2,86**	**2,78**	**2,69**	**2,61**	**2,52**	**2,42**
2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
3,31	**3,17**	**3,03**	**2,88**	**2,80**	**2,72**	**2,64**	**2,55**	**2,46**	**2,36**
2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
3,26	**3,12**	**2,98**	**2,83**	**2,75**	**2,67**	**2,58**	**2,50**	**2,40**	**2,31**
2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
3,21	**3,07**	**2,93**	**2,78**	**2,70**	**2,62**	**2,54**	**2,45**	**2,35**	**2,26**
2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
3,17	**3,03**	**2,89**	**2,74**	**2,66**	**2,58**	**2.49**	**2,40**	**2,31**	**2,21**
2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
3,13	**2,99**	**2,85**	**2,70**	**2,62**	**2,54**	**2,45**	**2,36**	**2,27**	**2,17**
2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
3,09	**2,96**	**2,81**	**2,66**	**2,58**	**2,50**	**2,42**	**2,33**	**2,23**	**2,13**
2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
3,06	**2,93**	**2,78**	**2,63**	**2,55**	**2,47**	**2,38**	**2.29**	**2,20**	**2,10**
2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
3,03	**2,90**	**2,75**	**2,60**	**2,52**	**2,44**	**2,35**	**2,26**	**2,17**	**2,06**
2,18	2,10	2,03	1,94	1,90	1,85	1,81	1.75	1.70	1,64
3,00	**2,87**	**2,73**	**2,57**	**2,49**	**2,41**	**2,33**	**2,23**	**2,14**	**2,03**
2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
2,98	**2,84**	**2,70**	**2,55**	**2,47**	**2,39**	**2,30**	**2,21**	**2,11**	**2,01**
2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
2,80	**2,66**	**2,52**	**2,37**	**2,29**	**2,20**	**2,11**	**2,02**	**1,92**	**1,80**
1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
2,63	**2,50**	**2,35**	**2,20**	**2,12**	**2,03**	**1,94**	**1,84**	**1,73**	**1,60**
1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
2,47	**2,34**	**2,19**	**2,03**	**1,95**	**1,86**	**1,76**	**1,66**	**1,53**	**1,38**
1,83	1,74	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00
2,32	**2,18**	**2.04**	**1,88**	**1,79**	**1,70**	**1,59**	**1,47**	**1,32**	**1,00**

Источник: Pearson, E. S., & Hartley, H. O. (1966). Biometrika tables for statistician Vol. 1, 3rd ed.). London Cambridge University Press.

Примечание. Для обычного шрифта критические значения для $\alpha = 0{,}05$. Для жирного шрифта критические значения для $\alpha = 0{,}01$.

ПРИЛОЖЕНИЕ Е

Ответы к заданиям для повторения

В этом приложении даны ответы к приведенным в конце каждой главы заданиям на множественный выбор и некоторым упражнениям. В ответах к заданиям на множественный выбор в скобках указывается номер страницы, на которой излагается соответствующий материал. В руководстве для преподавателей приведен полный список ответов к упражнениям.

Глава 1. Научное мышление в психологии

Множественный выбор:

1.3) (10) 2.1) (7) 3.4) (24) 4.2) (22) 5.2) (8)

Упражнения:

1.1. Эмпирические вопросы.

1) Бог умер? Возможный эмпирический вопрос: насколько сильно сторонники разных религий верят в личного Бога, участвующего в их ежедневной жизни?
3) Добры ли люди по своей природе? Возможный эмпирический вопрос: снизится ли вероятность благотворительных пожертвований, если люди будут знать, что они останутся анонимными?
5) Что такое красота? Возможный эмпирический вопрос: различаются ли определения женской красоты 20-летних и 40-летних мужчин?

Глава 2. Этика психологических исследований

Множественный выбор:

1-2) (34) 2.3) (49) 3.4) (44) 4.1) (59) 5.3) (61)

Упражнения:

2.1. Научное мышление и мистификация.

1) В исследовании может изучаться отношение студентов, участвовавших в экспериментах с и без мистификации, к исследованиям и психологам вообще. Кроме выявления отношения студентов можно попросить указать место психологов среди других профессий на «шкале доверия».

2.2. Обнаружение этических проблем.

1) Это исследование представляет собой проблему с точки зрения стандартов 6.08 (ничего не подозревавшие участники могут подать в суд за вмешательство в частную жизнь и даже спровоцировать применение уголовного наказания за оскорбление или осведомленного —см. вставку 3.1), 6.11 (отсутствие согласия; участникам не сказали, что они могут выйти из эксперимента), 6.15 (мистификация — уровень «негативных эмоциональных переживаний» в исследовании может быть слишком высок), 6.17 (минимизация постороннего воздействия). Чтобы оправдать исследование в глазах Исследовательского совета, экспериментаторы сделают особый акцент на том, что очень важно понять, как стресс влияет на психологические системы, что ситуация существенно не отличалась от того, что обычно происходит в мужском туалете, что исследование включает необходимый дебрифинг, что гарантируется конфиденциальность, а полученная информация не будет использована без разрешения участников.

Глава 3. Разработка идей для психологических исследований

Множественный выбор:

1.3) (73) 2. 2) (77) 3.3) (84) 4. 3) (93) 5.1) (94)

Упражнения:

3.1. Что дальше?

2) Исследование Милгрэма: один из вариантов — поместить жертву в одну комнату с учителем. Гипотеза будет заключаться в том, что при более явных страданиях уровень подчинения снизится (Милгрэм проделал такой эксперимент и уровень подчинения действительно несколько снизился).

3.2 Создание операциональных определений.

1. Фрустрацию можно операционально определить как: а) подъем кровяного давления на 20% выше обычного уровня или б) количество времени, затраченного на попытки преодолеть преграду, мешающую ребенку получить желанную игрушку.
3. Тревожность можно операционально определить как: а) уровень мышечного напряжения плеч или б) самостоятельная оценка человеком своей тревожности перед началом экзамена по шкале от 1 до 10.

Глава 4. Измерения, выборки и обработка данных

Множественный выбор:

1.3) (111) 2. 2) (112) 3.2) (114) 4. 2) (124) 5. 4) (117)

Упражнения:

4.2. Шкалы измерений.

1) Используются категории — номинальная шкала;

3) ранжирование — порядковая шкала;
5) номинальная шкала, если предположить, что в исследовании оценивалось, оказывают люди помощь или нет (т. е. две категории); но может быть и шкала отношений, если измеряется, сколько времени занимает помощь.

4.3. H_0, H_1, ошибки 1-го и 2-го рода.

1) H_0: мужчины и женщины, участвующие в исследовании, одинаково способны обнаружить обман. H_1: женщины лучше определят обман, чем мужчины (обоснование: чтение эмоций по лицу — это не «мужское дело»).

Ошибка 1-го рода: женщины превосходят мужчин, хотя на самом деле истинных различий в способности обнаруживать обман не существует. *Ошибка 2-го рода:* в исследовании не обнаружено различий, но в действительности женщины превосходят мужчин в способности обнаруживать обман.

Глава 5. Введение в экспериментальные исследования

Множественный выбор:

1.2)(160) 2. 4)(166) 3.1)(151) 4. 4)(157) 5. 4)(165)

Упражнения:

5.1. Выявление переменных.

1. В исследовании использованы две независимые переменные. Первая — это курс; она принимает два значения: первокурсник и студент старших курсов. Данная переменная является субъективной. Вторая переменная — местоположение здания; ее значения — центр и периферия; она является управляемой. Также использованы две зависимые переменные: степень уверенности (интервальная шкала) и точность указания направления (шкала отношений).
3. Использована одна управляемая независимая переменная, которую можно обозначить как «упорядочение сигнал-еда». Она принимает три значения: а) сигнал включен, выключен, затем еда, б) сигнал включен, затем еда и в) еда, затем сигнал. Также использованы две зависимые переменные: сколько времени требуется для начала саливации и объем саливации. Обе оцениваются по шкале отношений.

5.2. Обнаружение осложнителя(ей).

1. Независимой переменной является тип клюшки (четыре значения по четырем видам клюшек), а зависимой — расстояние, на которое откатывается мяч для гольфа. Осложнители — последовательность клюшек (гольферы не должны использовать клюшки в одном и том же порядке) и лунки

(некоторые лунки соответствуют более длинным дистанциям, если они расположены с подветренной стороны).

Значения НП	ВП1	ВП2	ЗП
Клюшка 1	первая по порядку	первая лунка	расстояние
Клюшка 2	вторая по порядку	вторая лунка	расстояние
Клюшка 3	третья по порядку	третья лунка	расстояние
Клюшка 4	четвертая по порядку	четвертая лунка	расстояние

3) Независимая переменная — стратегия запоминания (или что-то подобное). Она принимает два значения: зрительные образы и повторение. Осложнители — тип слов (все слова должны быть либо конкретными, либо абстрактными) и способ их представления (все слова должны представляться либо визуально, либо аудиалью). Зависимая переменная — количество запомненных слов.

Значения НП	ВП1	ВП2	НП
Зрительные образы	конкретные существительные	визуальное представление	запоминание
Механическое повторение	абстрактные существительные	аудиальное представление	запоминание

5.3. Операциональные определения (вновь).

1) Независимой переменной является неопределенность ситуации, и в случае эксперимента с оказанием помощи ею можно управлять, разыграв несчастный случай (например, человек со стоном сползает вниз по стене) и варьируя интенсивность освещения, используя, например, яркий, средний и тусклый свет. Освещение можно операционально определить в терминах физических характеристик яркости. Зависимой переменной может быть оказание помощи, которое можно операционально определить как поведение, проявляемое, когда прохожий останавливается и предлагает помочь.

Глава 6. Проблемы контроля в экспериментальных исследованиях

Множественный выбор:

1.4)(180) 2.3)(183) 3.3)(183) 4.1)(196) 5.2)(201)

Упражнения:

6.1 Межсубъектные или внутрисубъектные планы?

1) В исследовании необходимо изучить как молодых, так и старых животных — межсубъектная переменная. В каждой группе у одних животных зрительная зона мозга будет поражена, а у других останется целой — еще одна межсубъектная переменная.

3) Исследование будет заключаться в сравнении процедуры воздействия с какой-либо другой процедурой или с отсутствием воздействия (контрольная группа). В обоих случаях исследование межсубъектное.

5) В исследовании сравнивается способность к решению проблем в группе и в одиночестве. Его можно провести с использованием межсубъектной переменной, но внутрисубъектный подход также возможен в данном случае и даже предпочтителен, так как позволяет провести прямое сравнение и определить, в какой ситуации конкретные испытуемые лучше справляются с заданием, в группе или в одиночестве.

Глава 7. Экспериментальный план I: однофакторные планы

Множественный выбор:

1.3) (229) 2. 4) (218) 3.3)(217) 4. 2) (227) 5.1) (231?)

Упражнения:

7.1. Определение вида экспериментального плана.

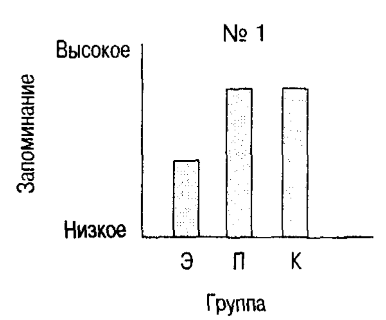
1) Независимой переменной является наличие или отсутствие у испытуемой булимии. Она является межсубъектной и субъективной. Зависимая переменная — выбор размера тела на рисунке. План однофакторный двухуровневый с неэквивалентными группами.

3) Независимая переменная — отсрочка вознаграждения — представляет собой межсубъектную управляемую переменную. Зависимая переменная — продолжительность занятия головоломками. План однофакторный многоуровневый с независимыми группами.

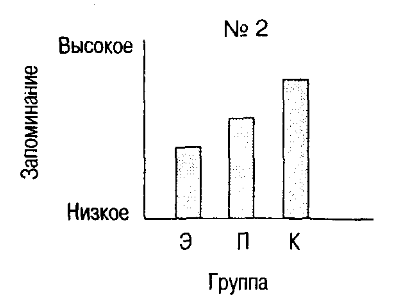
7.2. Результаты.

1) Независимая переменная является дискретной (тип группы в исследовании), поэтому необходимо использовать гистограмму.

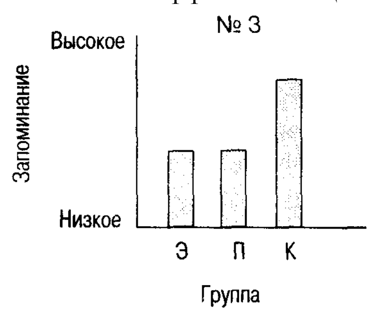
Результат А. Марихуана снижает запоминание, а ожидание испытуемым действия марихуаны не влияет на запоминание.



3) Результат В. Марихуана снижает запоминание, но ожидание испытуемым действия марихуаны также снижает запоминание.

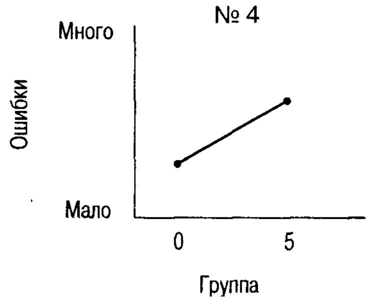


4) Результат С. Кажущееся неблагоприятное действие марихуаны на запоминание связано только с эффектом плацебо.

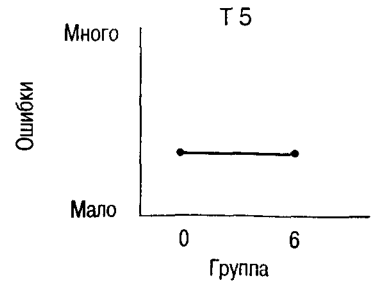


7.3. Независимая переменная является непрерывной (задержка подкрепления), поэтому лучше использовать линейный график (но гистограмма также возможна).

Результат А. Задержка подкрепления затрудняет обучение.



5) Результат В. Задержка подкрепления не влияет на обучение.



Глава 8. Экспериментальный план II: факторные планы

Множественный выбор:

1.3)(251) 2.4)(270) 3. 3)(246) 4.2)(262) 5.4)(263)

Упражнения:

8.1. Определение вида экспериментального плана.

1) Независимая переменная № 1: тип личности (А, В, промежуточный тип) (межсубъектная, субъективная).

   Независимая переменная № 2: когнитивное задание (видеоигра, ее отсутствие) (межсубъектная, управляемая).

   Зависимая переменная: точность оценки двухминутного интервала;

   План: факторный 2 χ 3 Р χ Е.

3) Независимая переменная № 1: участок кожи (10 различных участков) (внутрисубъектная, управляемая).

   Независимая переменная № 2: зрительные возможности (слепой, зрячий) (межсубъектная, субъективная).

   Независимая переменная № 3: время тестирования (утро, вечер) (межсубъектная, управляемая).

   Зависимая переменная: пороговые ощущения.

   План: смешанный факторный план 2х2х10РхЕ.

5) Независимая переменная № 1: цвет стимула (цветной, ч/б) (внутрисубъектная, управляемая).

   Независимая переменная № 2: расстояние до стимула (10 футов, 20 футов) (внутрисубъектная, управляемая).

   Зависимая переменная: ошибка.

   План: факторный план 2 χ 2 с повторяемыми измерениями.

Глава 9. Корреляционные исследования

Множественный выбор:

1.3)(285) 2.4)(296) 3.1)(306) 4.2)(301) 5.4)(294)

Упражнения: >

9.1. Интерпретация корреляций.

1) Возможно, что авторитарные матери не дают возможности своим детям выработать самостоятельность, и поэтому дети становятся застенчивыми (А > В). Но также возможно, что застенчивые от природы дети требуют большего контроля со стороны матерей (В > А).

3) Возможно, что книги из домашней библиотеки, прочитанные детьми, сделали их более умными, а следовательно, более успевающими студентами

(A > B). Возможно, что третья переменная — отношение родителей к знанию — привела к тому, что: а) в доме стало много книг и б) их дети стали хорошими студентами (C > A и B).

9.3. Интерпретация графиков рассеяния.

1) +0,50 для пирсонова r — это относительно высокое значение, что говорит о некотором сходстве познавательных процессов последовательной и одновременной обработки информации.

3) Данная корреляция говорит об отсутствии взаимосвязи между интеллектом и депрессией.

Глава 10. Квазиэкспериментальные планы и прикладные исследования

Множественный выбор:

1.1) (317) 2.3) (322) 3. 3) (323) 4. 3) (336) 5. 4) (340)

Упражнения:

10.1. Обнаружение факторов, угрожающих внутренней валидности.

1) Когда контрольная группа не используется, возможны несколько альтернативных объяснений. Поскольку 60% значительно ниже обычных 75% (т. е. это точка экстремума), регрессия является наиболее вероятной угрозой внутренней валидности. Также возможно влияние истории и отбора, а взросление менее вероятно, так как в исследовании участвуют две разные группы студентов.

3) Здесь наиболее вероятной является проблема отбора. Возможно, что между женщинами, согласившимися участвовать в исследовании, и случайно отобранными женщинами имеются значимые различия (например, они охотнее пробуют что-то новое).

5) Тридцать студентов начали изучать курс, но только 18 закончили его. Проблема заключается в истощении. Она проявляется в том, что группа, дошедшая до конца курса, отличается от группы, начавшей обучение (например, студенты более настойчивы). Эффективность методики обучения может быть низкой для тех студентов, которым необходимо более непосредственное руководство преподавателя.

Глава 11. Планы с малым *N*

Множественный выбор:

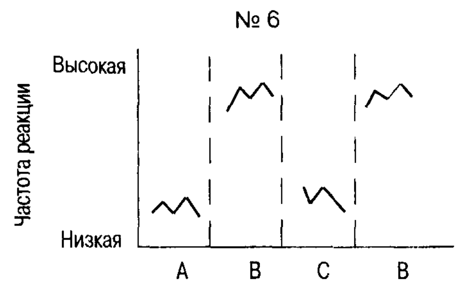
1.2) (353) 2.1) (357) 3. 3) (370) 4.1) (378) 5. 4) (369)

Упражнения:

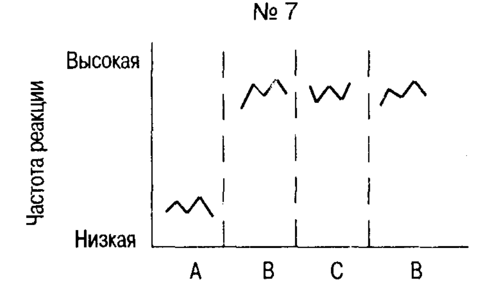
11.2. Гипотетические результаты прикладного анализа поведения.

1) A-B-C-B.

2) Подкрепление действует, но только если оно появляется при наличии определенного поведения.



3) Подкрепление действует вне зависимости от условий его появления.



Глава 12. Описательные методы исследований

Множественный выбор:

1.1) (390) 2.3) (393) 3.1) (398) 4.3) (410) 5. 3) (409)

Упражнения:

12.1. Усовершенствование опросников.

1) Страдали ли Вы расстройством желудка в последнее время?

   а) «Расстройство желудка» можно определить более точно, но основной проблемой является неточность словосочетания «последнее время». Что оно означает: последнюю неделю? месяц?

   б) Сказав испытуемому, что «расстройство желудка» означает «тошноту, которая, по его мнению, должна вызвать рвоту, но не вызывает» Вы можете сформулировать вопрос следующим образом: «страдали ли Вы расстройством желудка в последнюю неделю? Если да, то как часто?»

3) Насколько молод, по Вашему мнению, средний курильщик?

   а) Слова «насколько молод» могут привести к тому, что опрашиваемые укажут меньший возраст, а некоторые могут даже решить, что спрашивают о том, в каком возрасте люди начинают курить.

   б) Лучше спросить: «Каков, по Вашему мнению, средний возраст типичного курильщика?»

12.3. Выбор описательного метода.

1) Поскольку такие организации являются в основном частными, наблюдение участника исследования поможет больше узнать, чем естественное наблюдение, но войти в группу может быть непросто. Эффективными могут быть подробные интервью, проводимые в ходе качественного исследования.

3) Размер выборки слишком велик для проведения интервью, поэтому лучше использовать письменный или, возможно, телефонный опрос.

5) Это исследование конкретного примера, вероятно, включающее подробные интервью с пассажирами.

Глоссарий

ANOVA *{ANalysis Of VAriance).* Наиболее общий заключительный метод статистического анализа результатов эксперимента, при котором зависимые переменные измеряются в шкалах интервалов или отношений.

А77-план *{aptitude-treatment interaction design).* План взаимодействия способностей и внешнего воздействия. Образует факторный план Р χ Е, который часто встречается в исследованиях, посвященных вопросам обучения, и отображает взаимодействие между переменной способностей (личностный фактор *—person)* и переменной воздействия (фактор среды — *environment).*

Абсолютный порог *{absolute threshold).* Интенсивность стимула, достаточная для его обнаружения.

Авторитет *{authority).* Способ познания, предложенный Пирсом, при котором мнение формируется на основании другого мнения, составленного человеком, считающимся экспертом.

Альтернативная гипотеза *{alternative hypothesis).* Гипотеза экспериментатора о результатах исследования **(Н,).**

Анализ затрат и эффективности *{cost-effectiveness analysis).* Способ оценки затрат на разработку, проведение и завершение исследования.

Анализ потребностей *{needs analysis).* Вид предварительной оценки программы, при которой определяется наличие потребности в данной программе.

Априори метод *{a priori method)/* Способ познания, предложенный Пирсом, при котором человек формирует мнение с помощью самостоятельных рассуждений или на основании мнения других людей, составленного путем рассуждений.

Архивное исследование *{archivalresearch).* Описательный метод анализа полученной информации для проверки выдвинутой гипотезы.

Асимметричный перенос *{asymmetric transfer).* Возникает, когда одна последовательность действий вызывает эффект переноса, отличный от эффекта, вызванного второй, противоположной ей последовательностью.

База выборки *{sample frame).* Список людей, из которых делается выборка. При групповой выборке — список групп, из которых делается выборка.

Базовое исследование *{basic research).* Исследование, направленное на описание, предсказание и объяснение основных законов поведения.

Базовый уровень *{baseline).* Начальная стадия плана с малым *N,* в которой подлежащее изменению поведение предварительно наблюдается для определения обычной частоты реакции.

Блоковая рандомизация *{block randomization).* Процедура, используемая для осуществления случайного выбора и обеспечения равного количества испытуемых для каждого эксперимента в пределах одного исследования. Каждый эксперимент должен быть проведен, прежде чем какой-либо из них будет проводиться повторно с вновь выбранным испытуемым. Испытуемые выбираются случай-

ным образом. В качестве уравновешивающей процедуры блоковая рандомизация используется также в случае внутрисубъектных схем. Если испытуемые участвуют в эксперименте при каждом наборе условий более чем один раз, то они должны пройти все условия, прежде чем повторно проходить одно их них.

Боязнь оценки *{evaluation apprehension).* Особый вид тревожности, возникающий у испытуемых, приводящий к тому, что они начинают вести себя так, чтобы получить положительную оценку исследователя.

Валидность *{validity).* В общем смысле — степень действительного измерения *X,* а не *Y,* при измерении *X* (т. е. при валидном измерении интеллекта измеряется интеллект, а не что-то другое).

Величина эффекта *{effect size).* Количество воздействия, оказываемого одной переменной на другую; количество дисперсии зависимой переменной, которое может быть объяснено независимой переменной.

Взаимная корреляция *{partial con-elation).* Многомерная статистическая процедура для оценки эффекта, вызываемого третьей переменной. Если уровень корреляции *X* и У остается высоким после появлении третьего фактора Z, то Z считается третьей переменной.

Взаимодействие *{interaction).* В факторной схеме — присутствует, если результат воздействия одной независимой переменной зависит от значения другой независимой переменной.

Взросление *{maturation).* Угроза внутренней валидности исследования, возникающая, если участники переживают изменения, связанные со взрослением в ходе проведения исследования, а не с действием независимой переменной.

Внешнее измерение *{unobtrusive measure).* Измерение характеристик поведения, когда участники не знают о наличии наблюдения за ними.

Внешняя валидность *{external validity).* Возможность обобщения выводов исследования на других людей, условия и время.

Внутренняя валидность *{internal validity).* Степень независимости исследования от методологических ошибок, в особенности от смешения факторов.

Внутрисубъектный план *{within-subjects design).* Любой экспериментальный план, при котором для одних и тех же испытуемых создаются разные условия исследования. Также используется название «план с повторяющимися измерениями».

Возможность открытия *{discoverability).* Научное предположение о том, что применение научных методов позволяет находить причины событий.

Временная выборка *{time sampling).* Процедура, используемая при проведении наблюдений, заключающаяся в фиксировании поведения только в определенные моменты времени (например, каждые 10 минут).

Выборка *{sample).* Некоторая часть или подмножество популяции.

Гипотеза *{hypothesis).* Научное предположение о взаимосвязи некоторых переменных, проверяемое эмпирически.

Гистограмма *{histogram).* Столбиковая диаграмма частотного распределения.

Двойной слепой метод *{double blind)*. Контрольная процедура, уменьшающая искажения, при которой ни испытуемому, ни лицу, проводящему эксперимент, не известны условия проведения эксперимента. Часто применяется при изучении действия медикаментов.

Двумерный анализ *{bivariate analysis)*. Статистический анализ, выявляющий отношения двух переменных.

Дебрифинг *{debriefing)*. Общее собрание по окончании эксперимента, на котором исследователь объясняет истинные цели эксперимента, отвечает на вопросы участников и устраняет дискомфорт, возникший в ходе исследования.

Дедукция *{deduction)*. Доказательство от общего к частному. Применяется в науке при выведения из общей теории гипотезы конкретного исследования.

Демистификация *{dehoaxing)*. Часть дебрифинга, в которой участникам объясняется истинная цель эксперимента.

Десенсибилизация *{desensitizing)*. Часть дебрифинга, в которой экспериментатор устраняет нарушения нормального психического состояния или душевного равновесия, появившиеся в результате эксперимента.

Детерминизм *{determinism)*. Научное предположение о том, что любое событие имеет причину.

Диаграмма рассеивания *{scatterplot)*. Схема, отображающая отношения, показанные корреляцией.

Дискретная переменная *{discrete variable)*. Переменная, каждое значение которой представляет определенную категорию, отличную от другой категории (например, женщины и мужчины).

Дисперсия *{error variance)*. Несистематическое отклонение оценок, вызванное случайными факторами или индивидуальными различиями.

Дифференциальный порог *{difference threshold)*. Различие интенсивности стимулов 1 и 2, достаточное для дифференциации этих стимулов.

Единичное свидетельство *{anecdotal evidence)*. Единичный случай, иллюстрирующий некий феномен. Построение теории на основании единичного свидетельства, как в псевдонауке, приводит к ошибочным выводам.

Естественное наблюдение *{naturalistic obsewation)*. Описательный метод исследования, заключающийся в изучении поведения людей или животных в повседневной жизни.

Жизненный реализм *{mundane realism)*. Степень отражения экспериментом реальной жизни. Считается менее важным, чем экспериментальный реализм.

Зависимая переменная *{dependent variable)*. Поведение, расцениваемое как появившееся в результате эксперимента.

Заинтересованные лица *{stakeholders)*. При оценке программы — люди, связанные с программой и материально заинтересованные в ее проведении — заказчики, персонал и руководители программы.

Заключительное тестирование *{posttest)*. Измерение, проводящееся на заключительной стадии исследования, после того как участники (кроме членов кон-

трольной группы) подверглись экспериментальному воздействию. Полученные результаты сравниваются с результатами доэкспериментального исследования.

Закон *{law).* Постоянная предсказуемая связь событий.

Закрытый вопрос *{closed question).* Тип вопроса, применяемого при проведении опросов, на который можно ответить «да» или «нет» или обозначить ответ точкой на некоторой шкале.

Значение альфа *{alpha level).* Вероятность совершения ошибки 1-го рода; уровень значимости.

Индивидуальная валидность *{individual-subject validity).* Степень соответствия общих результатов исследования поведению отдельных участников.

Индукция *{induction).* Доказательство от частного к общему. Используется при изучении результатов исследования для доказательства или опровержения теории.

Инструктивная переменная *{instructional variable).* Независимая переменная, определяющая разные способы выполнения заданий для разных групп испытуемых.

Интервальная шкала *{interval scale).* Шкала измерений с равными делениями, по которым отсчитывается количество. Нулевая отметка является одной из многих и не означает отсутствия измеряемого феномена.

Интроспекция *{introspection).* Метод, широко применяемый на начальном этапе развития психологии, при котором испытуемый выполняет задания, а затем описывает явления, возникающие в сознании в процессе их выполнения.

Информационный подход *{data-driven).* Психологический исследовательский подход, заключающийся в том, что выводы о поведении делаются на основании информации, полученной научными методами.

ИС См. **исследовательский совет.**

Искажение, вызванное испытуемыми *{participant bias).* Возникает, когда на поведение участников исследования влияет их мнение о том, как они должны вести себя в ходе эксперимента.

Искажение, вызванное наблюдателем *{observer bias).* Возникает, если предвзятое мнение исследователя влияет на сущность наблюдений.

Искажение, вызванное отсутствием ответа *{nonresponse bias).* При проведении опроса — возникает, если люди, вернувшие опросники, имеют характерные отличия (например, в политической установке) от тех, кто не отвечал на вопросы.

Искажение, вызванное социальным ожиданием *{social desirability bias).* При проведении опроса — искажение ответа вследствии желания произвести наилучшее впечатление.

Искажение, вызванное экспериментатором *{experimenter bias).* Возникает, если ожидания экспериментатора влияют на результаты исследования.

Искажение по подтверждению *{confirmation bias).* Когнитивное искажение, при котором события, подтверждающие сложившееся мнение, воспринимаются и запоминаются с большей готовностью. Опровергающие мнение события игнорируются или забываются.

Исследование конкретного примера *{case study).* Описательный метод, при котором проводится всесторонний анализ отдельного индивидуума, редкого события или события, четко отражающего некоторый феномен.

Исследование методом поперечных **срезов** *{cross-sectional study).* В психологии развития — план, в котором возраст является независимой переменной. При этом тестируется несколько групп испытуемых, в каждой группе — люди одного возраста.

Исследовательская группа *{research team).* Группа исследователей (преподавателей и студентов), работающих вместе над какой-либо проблемой.

Исследовательский совет *{IRB — Institutional Review Board).* Университетская комиссия, занимающаяся изучением предложений о проведении научных исследований с целью оценки защиты прав участников. Существует во всех вузах, имеющих правительственное финансирование исследований.

История *{history).* Угроза внутренней валидности исследования, вызванная появлением в ходе исследования некоторого исторического события, влияющего на его участников.

Истощение *{attrition).* Нарушение внутренней валидности исследования, возникающая, когда испытуемые не могут продолжать участвовать в исследованиях. В основном это касается длительных исследований, при которых высока вероятность замены.

Качественное исследование *{qualitative research).* Вид исследования, заключающийся в нарративном анализе информации, полученной в ходе эксперимента. Включает изучение примеров, наблюдений и интервью.

Квазиэксперимент *{quasi-expenment).* Эксперимент, в котором не установлено влияние независимой переменной вследствие недостатков контроля над переменными в ходе исследования.

Кластерная выборка *{clustersample).* Вероятностная выборка, при которой случайным образом выбирается группа людей, имеющих нечто общее (например, студенты, посещающие лекции по истории), и тестируются все члены выбранной группы (т. е. все студенты трех из девяти групп, посещающих данные лекции).

Ключевой информатор *{key informant).* При оценке программы — член коллектива, осведомленный о потребностях данного коллектива.

Количественное исследование *{quantitative research).* Вид исследования, при котором результаты исследования представляются в цифровом виде, обычно в форме описательных и выведенных логически данных.

Конструкт *{construct).* Гипотетический фактор, например голод, который непосредственно не наблюдается, а предполагается на основании некоего поведения (еды) и является следствием определенных обстоятельств (24 часа без еды).

Конструктная валидность *{construct validity).* В измерениях — присутствует при точном оценивании некого гипотетического конструкта с помощью выбранной единицы измерений. Также показывает, является ли валидным сам конструкт. В исследованиях показывает валидность операциональных определений, использованных для независимых и зависимых переменных.

Контент-анализ *{content analysis).* Процедура, используемая в описательных исследованиях для категоризации исследуемого поведения (обычно вербального).

Контрольная группа *{control group).* Группа людей, не подвергавшихся воздействию, изучаемая для сравнения результатов.

Контрольная группа листа ожидания *{w aiting list control group).* Контрольная группа, члены которой еще не подвергались воздействию, но будут подвергнуты через некоторое время. Используется для подтверждения того, что члены экспериментальной и контрольной групп эквивалентны (например, и те и другие ожидают помощи в решении одной и той же проблемы).

Контрольная группа плацебо *{placebo control group).* Контрольная группа, некоторые члены которой уверены в том, что они подвергаются экспериментальному воздействию при отсутствии такового.

Контрольные таблицы поведения *{behavior checklists).* Таблицы с операциональными определениями, которые исследователи должны использовать при наблюдении.

Корреляционная матрица *{cross-laggedpanel correlation).* Применяется в корреляционных исследованиях, связанных с проблемой направленности. Если переменные *X* и *Y* измеряются в разное время и *X* предшествует У, то *X* может вызвать *Y,* а Уне может вызвать *X.*

Корреляция *{correlation).* См. **положительная корреляция** и **отрицательная корреляция.**

Коэффициент детерминации *{coefficient of determination).* Для двух коррелирующих факторов — степень изменчивости одного фактора, вызванная изменчивостью второго фактора. Находится возведением в квадрат пирсонова г.

Коэффициент корреляции *{coefficient of correlation),* см. **пирсоново *г {Pearson's* r).**

Креативное мышление *{creative thinking).* Процесс установления новых связей между идеями или событиями, на первый взгляд несвязанными.

Критериальная валидность *{criterion validity).* Валидность, состоящая в том, что наличие некоторого психологического критерия измерений дает возможность предсказать будущее поведение или этот критерий существенным образом связан с другим показателем.

Критериальная переменная *{criterion variable).* В регрессионном анализе — переменная, предсказываемая на основании предикторной переменной (например, школьные оценки предсказываются на основании оценок *SAT).*

Кумулятивное записывающее устройство *{cumulative recorder).* Аппарат для записи кумулятивной частоты реакции испытуемого при оперантном обусловливании.

Лабораторное исследование *{laboratory research).* Исследование, проводящееся в рамках научной лаборатории.

Латинский квадрат *{latine square).* Вид частичного позиционного уравнивания, при котором каждый набор условий эксперимента встречается в каждой позиции последовательности условий только один раз.

Линия регрессии *(regression line).* Линия, объединяющая точки на диаграмме рассеивания, дающая возможность делать прогнозы.

Лонгитюдное исследование *(longitudinal study).* В психологии развития — план, в котором возраст является независимой переменной, а одна и та же группа людей периодически исследуется на протяжении многих лет.

Медиана *(median).* Срединное значение набора данных — одинаковое количество значений располагаются выше и ниже медианы.

Межнаблюдательская надежность *(interobserver reliability).* Степень согласия исследователей друг с другом по поводу некоторого события.

Межсубъектный план *(between-subjects design).* Любой экспериментальный план, при котором для разных групп испытуемых создаются разные условия исследования.

Местоположение медианы *(median location).* Место в последовательности значений, где находится медиана.

Метод интервью *(interview suwey).* Вид опроса, при котором руководитель эксперимента лично беседует с участником. Этот метод обеспечивает более глубокое изучение при помощи последовательных вопросов и ответов.

Метод констант *(method of constant stimuli).* Психофизический метод, при котором стимулы с различной интенсивностью представлены в случайном порядке по одному (абсолютный порог) или попарно (дифференциальный порог).

Метод критических случаев *(critical incidents).* Метод, применяемый комиссией по этике, инспектирующей психологов, заключающийся в просьбе привести пример поведения, неэтического с точки зрения профессионального психолога.

Метод минимальных изменений *(method of limits).* Психофизический метод чередования возрастающих и убывающий серий стимулов. В экспериментах с возрастающей серией интенсивность стимула растет от подпороговой до поддающейся ощущению, с убывающей серией — уменьшается от надпороговой до неподдающейся ощущению.

Метод средней ошибки *(method of adjustment).* Психофизический метод, при котором испытуемый приспосабливается к интенсивности стимулов от слабой до нормальной их дифференциации.

Метод стебля и листа *(stem and leaf display).* Метод представления данных, заключающийся в объединении таблицы частотного распределения и гистограммы.

Мистификация *(deception).* Стратегия исследования, при которой участникам не сообщают подробностей проведения эксперимента с целью избежать появления наводящих признаков.

Множественная регрессия *(multiple regression).* Мультивариантный анализ, включающий критериальную переменную, а также две и более предикторных переменных, имеющих различный вес.

Мода *(mode).* Значение, наиболее часто появляющееся в наборе данных.

Мощность *(power).* Вероятность отрицания нулевой гипотезы (H_0), если она ложная.

Мультивариантный анализ *{multivariate analysis).* Статистический анализ, выявляющий отношения между тремя и более переменными.

Наблюдение участника исследования *{participant observation).* Описательный метод исследований, при котором изучается поведение людей в естественных условиях повседневной жизни, а исследователь является членом наблюдаемой группы.

Наводящий признак *{demandcharacteristic).* Особенность экспериментальной схемы или процедуры, повышающая вероятность того, что участники определят истинную цель эксперимента.

Надежность *{reliability).* Степень устойчивости измерений некоторого феномена и возможность их воспроизведения. Измерения с высокой надежностью содержат минимальное количество ошибок.

Наука *{science).* Способ познания, характеризующийся применением объективного эмпирического метода для поиска причин естественных явлений.

Независимая переменная *{independent variable).* Фактор, представляющий научный интерес, регулируемый исследователем (например, при внушении разной степени тревоги испытуемым) или подобранный заранее при отборе участников по степени проявления в них определенного качества (отбор двух групп людей с разным нормальным уровнем тревожности).

Нелинейный эффект *{nonlinear effect).* Результаты эксперимента, на графике отображающиеся в виде кривой. Возникает, только если независимая переменная принимает более двух значений.

Непрерывная переменная *{continuous variable).* Переменная, способная принимать бесконечное количество значений (например, дозировка лекарства).

Нерепрезентативная выборка *{biasedsample).* Выборка, не отражающая население в целом.

Номинальная шкала *{nominal scale).* Шкала измерений, в которой каждое деление означает не количество, а категорию, к которой можно отнести некоторое явление.

Нормальная кривая *{normalcurve).* Теоретическое частотное распределение популяции; кривая в форме колокола.

Нулевая гипотеза *{nullhypothesis).* Предположение о том, что между условиями проведения эксперимента нет действительных различий или что в исследовании по изучению корреляции не существует отношений между переменными (H_0).

Обратное позиционное уравнивание *{reverse counterbalancing).* Во внутрисубъектном плане при тестировании участников более одного раза при каждом наборе условий — испытуемые проходят одну серию экспериментов, а затем другую, расположенную в обратном порядке (т. е. A-B-C-C-B-A).

Общие собрания *{community forum).* При оценке программы — собрание всех участников для обсуждения необходимости определенной программы или особенностей ее проведения.

Объект *{subject).* Традиционное название участника исследования (человека или животного). В настоящее время людей, принимающих участие в исследовании в качестве испытуемых, называют участниками исследования, животных — объектами.

Объективность *{objectivity).* Предполагается ее наличие, если результаты наблюдения могут подтвердить несколько исследователей.

Объяснение *{explanation).* Задача науки, заключающаяся в поиске причин событий.

Ограничение диапазона *{restricting the range).* При изучении корреляции — использование ограниченного количества значений одной или обеих изучаемых переменных. Ограничение диапазона обычно снижает уровень корреляции.

Однофакторный многоуровневый план *{single-factor multilevel design).* План с одной независимой переменной, принимающей более одного значения.

Однофакторный план *{single-factor design).* Экспериментальный план с одной независимой переменной.

Оперантное обусловливание *{operant conditioning).* Вид научения, при котором поведение изменяется в зависимости от последствий — положительное последствие усиливает поведение, а отрицательное — ослабляет его.

Операциональное определение *{operationaldefinition).* Определение понятия или переменной путем точного описания операций, измерений или процедур.

Операционизм *{operationism).* Теория, разработанная Бриджманом в рамках философии науки, утверждающая, что значение понятия устанавливается рядом операций.

Описание *{description).* Задача науки, заключающаяся в точной классификации поведения или фиксировании последовательности внешних стимулов и соответствующих им поведенческих актов.

Описательная статистика *{descriptive statistics).* Краткое описание особенностей данных, полученных при исследовании некоторой выборки.

Оправдание усилий *{effort justification).* После больших затрат времени и сил на достижение некоторой цели человеку требуется убедить себя в том, что усилия были стоящими, даже если полученные результаты ниже, чем ожидалось.

Опрос *{survey).* Описательный метод, при котором испытуемые отвечают на серию вопросов или высказываются по поводу набора утверждений на заданную тему.

Опрос по телефону *{phone survey).* Метод опроса, при котором исследователь задает вопросы и получает ответы по телефону.

Осведомленное согласие *{informed consent).* Предоставление человеку достаточной информации об исследовании для принятия решения об участии в эксперименте.

Осложнитель *{confound).* Любая внешняя переменная, сходная с независимой переменной, которая приводит к альтернативному объяснению результатов.

Основной эффект *{main effect).* Показывает, существуют ли статистически значимые различия в значениях независимой переменной в факторном плане.

Открытый вопрос *{open-ended question).* Тип вопроса, использующегося при проведении опросов, предполагающий ответ, отличающийся от «да» или «нет».

Отрицательная корреляция *{negative correlation).* Отношение между переменными X и У, при котором высоким значениям X соответствуют низкие значения У, и наоборот.

Оценка программы *{program evaluation)*. Прикладное исследование, направленное на оценку программы от ее планирования до завершения.

Очевидная валидность *{face validity)*. Имеет место, когда один параметр измерения величины представляется более подходящим, чем другой (например, для измерения интеллекта способность к решению проблем имеет большую очевидную валидность, чем размер головы).

Ошибка 1-го рода *{type I error)*. Отказ от нулевой гипотезы, тогда как она верна, или признание наличия статистически значимого эффекта при его отсутствии.

Ошибка 2-го рода *{type lienor)*. Принятие нулевой гипотезы, когда она ошибочна, или признание отсутствия статистически значимого эффекта при его наличии.

Ошибка измерения *{measurement error)*. Ошибка, вызванная любым фактором, вводящим погрешность в измерение некоторой переменной.

Переменная задания *{task vanable)*. Вид независимой переменной, вводящейся, если каждый участник получает задание, отличающееся от остальных (например, лабиринты с различными уровнями сложности).

Переменная уравнивания *{matching vanable)*. Переменная, используемая для отбора участников исследования с эквивалентными группами.

Переоценка по запоминаемости *{availability heuristic)*. Когнитивное искажение, при котором люди переоценивают частоту появления ярких или запоминающихся событий.

Пилотажное исследование *(pilot study)*. Сбор данных на начальных этапах исследования. Проблемы, обнаруженные на этой стадии, позволяют усовершенствовать процедуру исследования и впоследствии избежать методологических ошибок.

Пирсоново *г (Pearson's r)*. Единица измерения корреляции между двумя переменными. Принимает значения от абсолютно негативной корреляции (-1) до абсолютно позитивной (+1). Если *г* = 0, то между переменными нет никаких отношений.

Письменный опрос *(written survey)*. Метод опроса, при котором исследователь разрабатывает письменный опросник, заполняемый затем участниками эксперимента.

Плагиат *(plagiarism)*. Сознательное присвоение чужих идей.

План А-А1-В-А1-В (А-А -В-А -В). План с малым *N* для оценки эффекта плацебо. А1 — условие, при котором применяется плацебо.

План А-В (А-В *design)*. План с малым *N,* в котором за базовым периодом (А) следует период воздействия (В).

План А-В-А (А-В-А *design)*. План с малым *N,* в котором за базовым периодом (А) следует период воздействия (В), а за ним — период, в котором воздействие заменяется противоположным или прекращается (второе А).

План А-В-А-В (А-В-А-В *design)*. То же, что и план **А-В-А,** но с последующим применением второго периода воздействия (второе В).

План А-В-С-В (А-В-С-В *design).* План с малым *N* для сравнения ситуативного подкрепления (В) с неситуативным (С), позволяющий исследователю разделить эффекты подкрепления и эффекты контингенции.

План с временной последовательностью *(time series design). Си.* **план с прерванной временной последовательностью.**

План с изменяющимся критерием *(changing criterion design).* План с малым *N,* в котором критерий получения подкрепления уточняется по ходу эксперимента. Используется для формирования поведения.

План с независимыми группами *(independentgroups design).* Межсубъектный план с регулируемой независимой переменной, предполагающий наличие не менее двух групп участников. Участники распределяются по группам случайным образом.

План с несколькими базовыми уровнями *(multiple baseline design).* План с малым *N,* при котором проводится воздействие с меняющимися интервалами (на разных уровнях, через разные интервалы) для изменения: а) поведения нескольких человек, б) нескольких вариантов поведения одного человека или в) поведения одного человека в разных условиях.

План с неэквивалентной контрольной группой *(nonequivalent control group design).* Квазиэкспериментальный план, в котором участники не могут быть случайным образом распределены на экспериментальную и контрольную группы.

План с неэквивалентными группами *(nonequivalentgroups design).* Межсубъектный план с субъектной переменной, включающий не менее двух групп участников; также план, при котором создаются неэквивалентные группы.

План с отменой *(withdrawal design).* План с малым *N,* при котором воздействие имеет место некоторое время, а затем прекращается для проверки возможности возвращения частоты ответов к базовому уровню.

План с повторяющимися измерениями *{repeated-measures design).* То же, что внутрисубъектный план. Участники тестируются при каждом наборе экспериментальных условий.

План с прерванной временной последовательностью *(interrupted time series design).* Квазиэкспериментальный план оценки программы исследований или воздействия на испытуемых путем измерения параметров выполнения заданий несколько раз до и несколько раз после реализации программы или воздействия.

План, прерванной временной последовательностью и переключающими повторениями *(interrupted time series with switching replications).* План с прерванной временной последовательностью, при которой программа повторяется в другом месте и в другое время.

План с уравненными группами *(matched groups design).* Межсубъектный план с регулируемой независимой переменной и не менее чем двумя группами участников. Испытуемые, эквивалентные в отношении некоторой переменной, влияющей на исход исследования, распределяются по группам случайным образом.

План с чередующимся воздействием *(alternating treatments design).* План с малым *N*ДЛЯ сравнения двух и более способов воздействия, направленных на измене-

ние некоторого поведения. План реализуется в рамках одного исследования, на примере одних и тех же испытуемых.

Позиционное уравнивание *{counterbalancing)*. Для внутрисубъектной переменной любая процедура, компенсирующая эффект последовательности.

Полевое исследование *{field research)*. Исследование, проходящее вне научной лаборатории.

Полевой эксперимент *{field experiment)*. Эксперимент, проводящийся за пределами лаборатории. Более узкий термин, чем «полевое исследование».

Полное позиционное уравнивание *{complete counterbalancing)*. Возникает, когда при внутрисубъектном плане порядок чередования условий эксперимента изменяется всеми возможными способами.

Положительная корреляция *{positive conelation)* Отношение между переменными *X и Y,* при котором высокое значение *X* соответствует высокому значению *Y,* и наоборот.

Популяция *{population)*. Совокупность всех членов идентифицируемой группы.

Порядковая шкала *{ordinalscale)*. Измерительная шкала, деления которой служат для относительного упорядочивания явлений.

Посторонняя переменная *{extraneous variable)*. Любой неуправляемый фактор, не представляющий научного интереса, влияющий на результаты эксперимента.

Предварительная оценка *{formative evaluation)*. Вид оценки программы, при котором рассматривается соответствие хода выполнения программы запланированному варианту.

Предварительное тестирование *{pretest)*. Измерение, проводящееся в начале исследования, до того как участники (кроме членов контрольной группы) подверглись экспериментальному воздействию.

Предикторная переменная *{predictor variable)*. В регрессионном анализе — переменная, используемая для прогнозирования значений критериальной переменной (например, школьные оценки предсказываются на основании оценок *SAT)*.

Прикладное исследование *{applied research)*. Исследование, направленное на решение конкретных, не допускающих отлагательств проблем.

Прикладной анализ поведения *{applied behavior analysis)*. Исследование, использующее различные методы оценки эффективности процедур обусловливания, призванных вызвать изменение скорости реакции.

Проблема направленности *{directionality problem)*. В корреляционных исследованиях — при корреляции двух переменных *X* и *Y, X* вызывает *Y,* но *Y* также может вызвать *X.* В таком случае корреляция не дает достаточных оснований для выбора одного из вариантов.

Проблема самовыбора *{self-selection problem)*. При проведении опросов — если выборка состоит только из людей, добровольно вызвавшихся отвечать, то она может не быть репрезентативной.

Проблема третьей переменной *{third variable problem)*. Проблема, связанная с направлением выводов корреляционного исследования в определенное русло,

вызванным третьей переменной — любым неуправляемым фактором, подчеркивающим корреляцию переменных *XиY*.

Проверка по критерию Стьюдента для зависимых групп *(t test for dependent groups).* Дедуктивный статистический анализ, используемый при сравнении двух групп в плане с эквивалентными группами или с повторяющимися измерениями.

Проверка по критерию Стьюдента для независимых групп (i *testfor independent groups).* Дедуктивный статистический анализ, используемый при сравнении двух групп в плане с независимыми или неэквивалентными группами.

Проверка эффективности воздействия *{manipulation check).* При дебрифинге — выявление, насколько субъект знаком с истинными целями эксперимента; также проверка, вызывает ли систематическое воздействие на участников эксперимента предполагаемый эффект.

Прогнозирование *(predictions).* Задача науки, которая заключается в утверждении возникновения некоторого поведения. Утверждение имеет определенную степень вероятности.

Программный аудит *(program audit).* Изучение соответствия реализации программы запланированному варианту. Вид предварительной оценки.

Программы исследований *(programs of research).* Совокупность взаимосвязанных исследований, в которой результаты одного эксперимента приводят к другому эксперименту.

Продуктивность *(productivity).* Связана с количеством исследований, разработанных для проверки выдвинутой теории. Теории, приводящие к большому количеству исследований, считаются продуктивными.

Простая случайная выборка *(simple random sample).* Вероятностная выборка, при которой все члены популяции имеют равные шансы быть выбранными.

Простая теория *(parsimonious).* Теория, содержащая минимальное число конструктов и допущений, необходимых для правильного объяснения и предсказания некоторого явления.

Псевдонаука *(pseudoscience).* Исследования, опирающиеся на единичные свидетельства и слишком неопределенные для адекватной проверки.

Психофизика *(psychophysics).* Одна из областей исследований экспериментальной психологии; исследует отношения между физическим стимулом и его ощущением; изучает пороги ощущения.

Разброс *(range).* Для набора оценок — разница между оценками, имеющими наибольшее и наименьшее значения.

Разглашение *(leakage).* Сообщение предыдущими участниками эксперимента истинных целей исследования будущим испытуемым.

Расслоенная выборка *(stratifiedsample).* Случайная вероятностная выборка, в которой пропорционально представлены представляющие интерес подгруппы.

Расширение *(extension).* Повторение основного исследования при изменении некоторых факторов (например, введение дополнительного значения независимой переменной).

Расщепленная надежность *{split-half reliability).* Вид надежности, при которой половина элементов теста (например, с четными номерами) коррелирует с другой половиной.

Реактивность *{reactivity).* Наблюдается, когда на поведение участников исследования влияет знание того, что за ними наблюдают.

Регрессионный анализ *{regression analysis).* В корреляционных исследованиях — процесс, заключающийся в предсказании значения переменной *Y* по степени корреляции и значению переменной *X.*

Регрессия к среднему *{regression to the mean).* Если при тестировании получены очень высокие или низкие оценки, то при следующем тестировании оценки будут ближе к среднему. Регрессия к среднему может снизить внутреннюю валидность эксперимента, если доэкспериментальные оценки очень высокие или низкие, а послеэкспериментальные — ближе к среднему.

Репликация *{replication).* Повторение эксперимента. Точные повторения редки и обычно проводятся при наличии предполагаемых ошибок в основном эксперименте.

Репрезентативная выборка *{representative sample).* Выборка, характеристики которой соответствуют характеристикам популяции.

Ретестовая надежность *{test-retest reliability).* Вид надежности, при котором один тест проводится два раза и определяется корреляция между этими попытками.

Ригидность мнения *{belief perseverance).* Нежелание рассматривать никакие факты, противоречащие придерживаемой точке зрения. Близко к принципу **упорства,** предложенному Пирсом.

Рисунок *{figures).* В лабораторном отчете или описании исследования — графики, диаграммы, схемы, наброски приборов или фотографии.

Роль положительного субъекта *{good subject role).* Вид искажения, вызванного испытуемыми, при котором они пытаются угадать гипотезу исследования и подтвердить ее своим поведением.

Систематическая дисперсия *{systematic variance).* Изменчивость по причине, поддающейся определению, — либо вследствие систематического изменения независимой переменной, либо из-за неконтролируемого изменения результата.

Ситуационная переменная *{situational variable).* Внешняя переменная, вводимая, если испытуемые сталкиваются с различными внешними обстоятельствами (например, противопоставление большой и маленькой комнат при проведении исследования с большим количеством участников).

Случайное открытие *{serendipity).* Нахождение *X* при поиске *Y.*

Случайное распределение *{random assignment).* Наиболее общий метод создания эквивалентных групп при реализации межсубъектного плана. Все испытуемые имеют равную вероятность попасть в ту или иную группу.

Смешанный факторный план *{mixed factorial design).* Факторный план с межсубъектными и внутрисубъектными факторами.

Событийная выборка *(event sampling).* Процедура наблюдения, в которой рассматриваются только определенные варианты поведения, выявленные в конкретных условиях.

Согласие *(assent).* В кодексе *SRCD,* посвященном этике исследований, связанных с детьми, согласие означает готовность ребенка участвовать в исследовании.

Социальная валидность *(social validity).* Потенциальная возможность программы анализа поведения улучшить общество в процессе ее проведения или вследствие осознания ее ценности участниками исследования.

Среднее арифметическое *(mean).* Значение, полученное суммированием всех значений определенного набора данных и делением суммы на общее количество значений.

Стандартное отклонение *(standard deviation).* Отклонение набора значений от среднего; квадратный корень дисперсии.

Статистика вывода *(inferential statistics).* Обобщение выводов исследования выборки на население в целом.

Статистическая валидность *(statistical conclusion validity).* Степень точности статистического анализа и сделанных на его основании выводов.

Статистический детерминизм *(statistical determinism).* Предположение исследователей-психологов о том, что поведение можно предсказать с вероятностью выше вероятности случая.

Субъективная переменная *(subject variable).* Дополнительная переменная, которая выбирается экспериментатором и не изменяется в ходе эксперимента. Любая характеристика индивидуума, выбранного для проведения исследования (например, пол).

Суммарная оценка *(summative evaluation).* Вид оценки программы, проводящейся в конце программы с целью определить эффективность выполнения поставленных задач.

Сходящиеся операции *(converging operations).* Возникают, когда результаты нескольких различных исследований, по-разному определяющих термины, сходятся к общему выводу.

Таблица *(table).* В отчете об исследовании — представление данных или описание плана исследования в виде рядов и столбцов.

Тенденция *(trend).* Предсказуемый набор событий, происходящих в некоторый период; выявляется в исследованиях с временным планом.

Теория *(theory).* Набор утверждений, организующих и обобщающих информацию о некотором явлении, представляющий собой основу для прогнозов, достоверность которых проверяется затем эмпирически.

Тестирование *(testing).* Нарушение внутренней валидности вследствие влияния доэкспериментального исследования на результаты послеэкспериментального исследования (например, возникающего из-за привыкания испытуемых).

Удобная выборка *(convenience sample).* Простая выборка, при которой исследователь вызывает добровольцев из группы людей, соответствующих общим тре-

бованиям проведения исследования (например подростков). Применяется в большинстве психологических исследований, кроме тех, которые направлены на анализ потребностей населения.

Упорство *(tenacity).* Способ познания, описанный Пирсом, заключающийся в формировании искаженных представлений и нежелании изменять их перед лицом опровергающих их данных.

Управление *(control).* Задача науки, которая заключается в применении основных законов, открытых научными методами, для решения проблем.

Уравнивание *(matching).* Процедура, при которой у всех участников измеряется определенная характеристика («переменная отбора»), коррелирующая с зависимой переменной, а затем участники с одинаковыми значениями переменной отбора случайным образом распределяются по группам.

Участник исследования *(research participant).* Человек, принимающий участие в исследовании в качестве испытуемого.

Факторная матрица *(factorial matrix).* Таблица, описывающая факторную схему, в рядах и столбцах которой отображаются независимые переменные, значения этих переменных и общее число условий эксперимента (ячейки таблицы).

Факторный анализ *(factor analysis).* Многомерный анализ, выявляющий взаимосвязь многочисленных переменных. Переменные, высоко коррелирующие друг с другом, образуют «факторы».

Факторный план *(factorial design).* Экспериментальный план, имеющий более одной независимой переменной.

Факторный план *PxE(Px* *E factorial design).* Факторный план, включающий не менее одного субъектного фактора *(P = personal variable* —переменная индивидуума) и не менее одного управляемого фактора *(E = environmental variable* — переменная среды).

Фальсификация *(falsification).* Исследовательский подход, защищаемый Поппером, заключающийся в проверке теорий путем попытки их опровержения или замены.

Фальсификация данных *(falsifyingdata).* Обработка или подмена данных для получения желаемого результата.

Фокус-группа *(focusgroup).* Небольшая и относительно однородная группа людей, собранная для коллективного обсуждения некоторой темы или — при оценке программы — для выяснения необходимости определенной программы или особенностей ее проведения.

Фонд участников *(participantpool).* Группа студентов, которых попросили принять участие в начальном этапе исследования в качестве испытуемых. Историческое название — «фонд испытуемых».

Формирование *(shaping).* Оперантная процедура для выработки нового поведения, лежащая в основе плана с изменяющимся критерием. Поведение подкрепляется в процессе приближения к требуемой модели.

Хоторнский эффект *{Hawthorne effect)*. Предубеждение участников исследования, считающих, что участие в эксперименте повышает их значимость в глазах экспериментатора.

Частичное повторение *{panial replication)*. Повторение части некоторого исследования. Обычно проходит в рамках эксперимента, расширяющего результаты основного исследования.

Частичное позиционное уравнивание *{panial counterbalancing)*. Возникает, когда во внутрисубъектном плане использованы многие из возможных перестановок условий.

Частота реакции *{rate of response)*. Зависимая переменная, часто изучаемая исследователями школы Скиннера. Означает частоту появления некоторого поведения в единицу времени.

Частотное распределение *{frequency distribution)*. Таблица, отражающая количество появлений каждой оценки в наборе других оценок.

Шкала измерений *{measurementscales)*. Способ измерения явления. См. номинальная шкала, порядковая шкала, интервальная шкала и шкала отношений.

Шкала отношений *{ratio scale)*. Измерительная шкала с равными делениями, в которой цифры означают количество. Нулевая отметка означает отсутствие измеряемого феномена.

Эквивалентная контрольная группа *{yoked control group)*. Контрольная группа, члены которой подвергаются такому же воздействию, как и члены экспериментальной группы.

Экологическая валидность *{ecological validity)*. Рассматривается при изучении психологических феноменов повседневной жизни (например, способность запомнить место, где оставлены ключи).

Эксперимент *{experiment)*. Процедура исследования, при которой одни факторы остаются постоянными, другие изменяются и фиксируется некий результат.

Экспериментальная группа *{experimentalgroup)*. В исследованиях с определенной контрольной группой — группа людей, подвергшаяся определенному, требующему изучения воздействию.

Экспериментальный реализм *{experimental realism)*. Степень вовлеченности испытуемых в эксперимент. Более важно, чем жизненный реализм.

Эмпиризм *{empirism)*. Способ познания, основывающийся на непосредственном наблюденнии или опыте.

Эмпирический вопрос *{empirical question)*. Вопрос, ответ на который может быть получен с помощью объективного наблюдения.

Этика *{ethics)*. Набор принципов, предписывающих моральное поведение.

Эффект когорты *{cohort effect)*. Когорта — это группа людей, рожденных в одно время. Эффект когорты снижает внутреннюю валидность исследования методом поперечных срезов, так как между группами возникает различие из-за того, что члены этих групп выросли в разные исторические периоды.

Эффект отбора испытуемых *(subject selection effect)*. Нарушение внутренней валидности исследования, возникающая, если испытуемые не могут быть распределены по группам случайным образом, а следовательно, группы не являются эквивалентными.

Эффект передачи *(carryover effect)*. Вид эффекта последовательности, при котором устойчивые изменения в выполнении заданий возникают в результате выполнения конкретной последовательности условий, а не какой-либо другой последовательности.

Эффект порядка *(order effect)*. См. **эффект последовательности.**

Эффект последовательности *(sequence effect)*. При реализации внутрисубъектного плана — влияние опыта, полученного при одних условиях проведения эксперимента, на последующее выполнение заданий при других условиях.

Эффект потолка *(ceiling effect)*. Появляется, когда результаты, полученные при различных условиях, близки к максимальным для данной шкалы и создается впечатление, что они были получены при одинаковых условиях.

Эффект прогрессии *(progressive effect)*. Для внутрисубъектного плана — эффект последовательности, при котором изменение состояния испытуемого происходит равномерно от одной попытки к другой (например, накапливается усталость).

Литература

Abramovitz, C. V., Abramovitz, S. I., Roback, H. B., & Jackson, M. C (1974). Differential effectiveness of directive and nondirective group therapies as a function of client internal-external control. *Journal ofConsulting and Clinical Psychology, 14,*849-853.

Abramson, L.Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: A critique and reformulation. *Journal of Abnormal Psychology, 87,* 49-74.

Adair, J. G. (1973). *The human subject.The social psychology of the psychological experiment.* Boston: Little, Brown.

Adams, J. S. (1965). *Inequity in social exchange.* In L. Berkowitz (Ed.), Advances in experimental social psychology (pp. 267-299). New York: Academic Press.

Adler, T. (1992, September). Debate: Control groups — bad for cancer patients? *APA Monitor, 23,* 34.

Adler, T. (1992, November). Trashing a laboratory is now a federal offense. *APA Monitor, 23,* 14.

American Heritage Dictionary. (1971). New York: American Heritage.

American Psychological Association. (1953). *Ethical standards of psychologists.* Washington, DC; Author.

American Psychological Association. (1973). *Ethical principles in the conduct of research with human participants.* Washington, DC: Author.

American Psychological Association. (1982). *Ethical principles in the conduct of research with human participants.* Washington, DC: Author.

American Psychological Association. (1985). *Guidelines for ethical conduct in the care and use of animals.* Washington, DC: Author.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597-1611.

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.) J.Washington, DC: Author.

Anderson, C A., Lepper, M. R., & Ross, L. (1980). Perseverance of social beliefs: The role of explanation in the persistence of discredited *information.Journal of Personality and Social Psychology, 39,* 1037-1049.

Anderson, C A., Lindsay, J.J.,& Bushman, B.J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science, 8,* 3-9.

Anderson, J. E. (1926). Proceedings ofthe thirty-fourth annual meeting ofthe American Psycho-logical Association. *Psychological Bulletin, 23,* 113-174.

Anderson, K.J. (1990). Arousal and the inverted-U hypothesis: A critique of Neiss's «reconceptualizing arousal.» *Psychological Bulletin, 107,* 96-100.

Animal rights activity increases: Threats made against behavioral scientists. (1997, January/Feb-ruary). *Science Agenda,1'0(1),* 1,4.

Anonymous Advertisement. (1881, October). *Phrenological Journal, 73,* old series, 3-4.

Aronson, E. (1999a). *The social animal* (8th ed.). New York: W. H. Freeman.

Aronson, E. (1999b). Dissonance, hypocrisy, and the self-concept. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 103-126). Washington, DC: American Psychological Association.

Aronson, E., Fried, C, & Stone, J. (1991). Overcoming denial and increasing the intention to use condoms through the induction of hypocrisy. *American Journal of Public Health, 81,* 1636-1638.

Asch, S. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs, 70,* (Whole No. 416).

Atkinson, J. W., & Feather, N. T. (1966). *A theory of achievement motivation.* New York: Wiley.

Award for distinguished scientific contributions: Eliot Aronson (1999). *American Psychologist, 54,* 873-884.

Babkin, B. P. (1949). *Pavlov: A biography.* Chicago: University of Chicago Press.

Bahrick, H. P. (1984). Semantic memory content in permastore: 50 years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1-29.

Bakan, D. (1966). The influence of phrenology on American *psychology .Journal of the History of the Behavioral Sciences, 2,* 200-220.

Bakeman, R., & Brownlee, J. R. (1980). The strategic use of parallel play: A sequential analysis. *Child Development, 51,* 873-878.

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology, 66,* 3-11.

Barber, T. X. (1976). *Pitfalls in human research.* New York: Pergamon Press.

Barlow, D. H., & Hersen, M. (1984). *Single case experimentaldesigns: Strategies forstudying behavior change* (2nd ed.). New York: Pergamon Press.

Barnett, M. A., & McCoy, S.J. (1989).The relation of distressful childhood experiences and empathy in college *undergraduates.Journal of Genetic Psychology, 150,*417-426.

Baron, A. (1990). Experimental designs. *The Behavior Analyst, 13,* 167-171.

Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's «Behavioral study of obedience.» American Psychologist, 19, 421-423.

baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist, 40,* 165-174.

Bavelas, J. B. (1978). *Personality: Current theory and research.* Monterey, C A: Brooks/ Cole.

Beauchamp, t. L., & Childress, J. F. (1979). *Principles of biomedical ethics.* New York: Oxford University Press.

Behr, W. A. (1992). Ph.D. envy: A psychoanalytic case study. *Clinical Social Work Journal, 20,* 99-113.

Beisecker, T. (1988). Misusing survey research data: How not to justify demoting Christine Craft. *Forensic Reports, 1,* 15-33.

Benedict, J., & Stoloff, M. (1991). Animal laboratory facilities at «America's Best» undergraduate colleges. *American Psychologist, 46,* 535-536.

Benjamin, L. T. Jr., Cavell, T. A., & Shallen-Berger, W. R. (1984). Staying with initial answers on objective tests: Is it a myth?. *Teaching of Psychology, 11,* 133-141.

Berry, T. D., & Geller, E. S. (1991). A single-subject approach to evaluating vehicle safety belt reminders: Back to *bancs. Journal of Applied Behavior Analysis, 24,* 13-22.

Bertera, R. L. (1990). Planning and implementing health promotion in the workplace: A case study of the Du Pont Company experience. *Health Education Quarterly, 17,* 307-327.

Beyerstein, B. (1993, July). Subliminal self-help tapes: Promises, promises... *Rational Enquirer, 6(1).*

Blagrove, M. (1996). Effects of length of sleep deprivation on interrogative suggestibility. *Journal of Experimental Psychology: Applied, 2,* 48-59.

Blakemore, C, & Cooper, G. F. (1970). Development of the brain depends on the visual environment. *Nature, 228,* 477-478.

Blumberg, M., & Pringle, C. D. (1983). How control groups can cause loss of control in action research: The case of Rushton coal *mine. Journal of Applied Behavioral Science, 19,* 409-425.

Boesch-Achermann, H., & Boesch, C (1993). Tool use in wild chimpanzees: New light from dark forests. *Current Directions in Psychological Science, 2,* 18-21.

Boorstin, D.J. (1985). *The discoverers.* New York: Vintage Books.

Boring, E. G. (1950). *A history of experimental psychology* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Bouchard, T. *J.,* Lykken, D. T., McGue, M., Segal N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250,* 223-228.

Bouchard, T. J., & McGuE, M. (1981). Familial studies of intelligence: A review. *Science, 212,*1055-1059.

Brady, J. V. (1958, April). Ulcers in «executive» monkeys. *Scientific American, 199,*95-100.

Brady, J. V., Porter, R. W., Conrad, D. G., & Mason, J. W. (1958). Avoidance behavior and the development of gastroduodenal ulcers. *Journal of the Experimental Analysis of Behavior, 1,* 69-72.

Bramel, D., & Friend, R. (1981). Hawthorne, the myth of the docile worker, and class bias in psychology. *American Psychologist, 36,* 867-878.

Bransford, J. D„ & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717-726.

Rantjes, M., & Bouma, A. (1991). Qualitative analysis of the drawings of Alzheimer's patients. *The Clinical Neuropsychologist, 5,* 41-52.

BrennanJ. F. (1991). *History and systems of psychology:* Englewood Cliffs, NJ: Prentice-Hall.

Bretz, R. D., Jr., & Thomas, S. L. (1992). Perceived equity, motivation, and final-offer arbitration in major league *baseball. Journal of Applied Psychology, 77,* 280-287.

Bridgman, P. W. (1927). *The logic of modern physics.* New York: Macmillan.

Broadbent, D. E. (1958). *Perception and communication.* New York: Pergamon Press.

Brown, M. F. (1992). Does a cognitive map guide choices in the radial-arm maze?*Joumal of Experimental Psychology: Animal Behavior Processes, 18,* 56-66.

Bryan, J. H., & Test, M. A. (1967). Models and helping: Naturalistic studies in aiding behavior. *Journal of Personality and Social Psychology, 6,* 400-407.

Burks, B. S., Jensen, D. W., & Terman, L. (1930). *Genetic studies of genius, Vol. 3.The promise of youth: Follow-up studies of a thousand gifted children.* Stanford, C A: Stanford University Press.

Burnham, J. C (1972).Thorndike's puzzle *hoxes.Joumalof'the History of the Behavioral Sciences, 8,* 159-167.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10,* 243-248.

Byrne, G. (1988, October 7). Breuning pleads guilty. *Science, 242,* 27-28.

Campbell, D. T. (1969). Reforms as experiments. *American Psychologist, 24,* 409-429.

Campbell, D. T., & Erlebacher, A. (1970). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), *Compensatory education: A national debate* (pp. 185-210). New York: Brunner-Mazel.

Campbell, D. T., & Ross, H. L. (1968).The Connecticut crackdown on speeding: Time series data in quasi-experimental analysis. *Law and Society Review, 3,* 33-53.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand-McNally.

Carnap, R. (1966). *An introduction to the philosophy of science.* New York: Basic Books.

Carello, C, Anderson, K. L., & Kunkler-Peck, A.J. (1998). Perception of object length by sound. *Psychological Science, 9,* 211-214.

Carr, H. A., & Watson, J. B. (1908). Orientation in the white rat. *Journal of Comparative Neurology and Psychology, 18,* 27-44.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15,* 373-380.

Cattell, J. M. (1895). Proceedings of the third annual meeting of the American Psychological Association. *Psychological Review, 2,* 149-172.

Christensen, L. (1988). Deception in psychological research: When is its use justified? *Personality and Social Psychology Bulletin, 14,* 664-675.

Christenson, C V. (1971). *Kinsey.A biography.* Bloomington: Indiana University Press.

Cicirelli, V. G. (1984).The misinterpretation of the Westinghouse study: A reply to Zigler and Berman. *American Psychologist, 39,* 915-916.

Cicirelli, V. G. (1993). Head Start evaluation. *APS Observer, 6(1),* 32.

Cicirelli, V. G., Cooper, W. H., & Granger, R. L. (1969).The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development. *Westinghouse Learning Corporation, OEO Contract B89-4536.*

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, S., Tykell, D. A., & Smith, A. P. (1993). Negative life events, perceived stress, negative affect, and susceptibility to the common cold. *Joumal of Personality and Social Psychology, 64,* 131-140.

Coile, D. C, & Miller, N. E. (1984). How radical animal activists try to mislead humane people. *American Psychologist, 39,* 700-701.

Converse, J. M., & Presser, S. (1986). *Suwey questions: Hand crafting the standardized questionnaire.* Newbury Park, CA: Sage.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimental design and analysis issues for field settings.* Chicago: Rand-McNally.

Coulter, X. (1986). Academic value of research participation by undergraduates. *American Psychologist, 41,317.*

Crawford, H. J., Kitner-Triolo, M., Clarke, S. W., & Olesko, B. (1992). Transient positive and negative experiences accompanying stage *hypnosis.fournalofAbnormal Psychology, 101,663-667.*

Cronbach, L.J. (1957).The two disciplines of scientific psychology. *American Psychologist, 12,* 671-684.

Cronbach, L. J., Hastorf, A. H., Hilgard, E. R., & Maccoby, E. E. (1990). Robert R. Sears (1908-1989). *American Psychologist, 45,* 663-664.

Dallenbach, K. M. (1913). The measurement of attention. *American Journal of Psychology, 24,* 465-507.

Danziger, K. (1985). The origins ofthe psychological experiment as a social institution. *American Psychologist, 40,* 133-140.

Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of *responsibility. Journal of Personality and Social Psychology, 8,* 377-383.

Dekay, M. L., & McClelland, G. H. (1996). Probability and utility components of endangered species preservation programs *Journal of Experimental Psychology: Applied, 2,* 60-83.

DeLuca, R. V., & Holborn, S. W. (1992). Effects of a variable-ratio schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25,* 671-679.

Department of Health and Human Services. (1983). Federal regulations for the protection of human research subjects. In L. A. Peplau, D. O. Sears, S. E. Taylor, & J. L. Freedman (Eds.), *Readings in social psychology* (2nd ed.). Englewood Cliffs, NT: Prentice-Hall.

Dermer, M. L., & Hoch, T. A. (1999). Improving descriptions of single-subject experiments in research texts written for undergraduates. *Psychological Record, 49,*49-66.

Dewsbury, D. A. (1990). Early interactions between animal psychologists and animal activists and the founding of the APA committee on precautions in animal experimentation. *American Psychologist, 45,* 315-327.

Diener, E., & Crandall, R. (197'8). *Ethics in social and behavioral research.* Chicago: The University of Chicago Press.

Domjan, M., & Purdy, J. E. (1995). Animal research in psychology: More than meets the eye of the general psychology student. *American Psychologist, 50,* 496-503.

Donnerstein, E. (1980). Aggressive erotica and violence against women. *Journal of Personality and Social Psychology, 39,* 269-277.

Dressier, F. B. (1893). On the pressure sense of the drum of the ear and «facial-vision.» *American Journal of Psychology, 5,* 344-350.

Dror, I. E., Kosslyn, S. M., & Waag, W. L. (1993). Visual-spatial abilities of *pilots.Journal of Applied Psychology,* 78,763-773.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology, 30,* 510-517.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C A. Bussenius, Trans.). New York: Dover. (Original work published 1885).

Egeland, B. (1975). Effects of errorless training on teaching children to discriminate letters of the alphabet. *Journal of Applied Psychology, 60,* 533-536.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, politics, and marriage.* New York: W.W. Norton.

Elkins, I. J., Cromwell, R. L., & Asarnow, R. F. (1992). Span of apprehension in schizophrenic patients as a function of distractor masking and laterality.*Journal of Abnormal Psychology, 101,* 53-60.

Elmes, D. G., Kantowitz, B. H., & Roediger, H. L., III. (1992). *Research methods in psychology* (4th ed). St. Paul, MN: West.

Emde, R. N., Plomin, R., Robinson, J., Corley, R., Defries, J., Fulker, D. W., Reznick, J. S., Campos, J., Kagan, J., & Zahn-Waxler, C (1992). Temperament, emotion, and cognition at fourteen months: The MacArthur longitudinal twin study. *Child Development, 63,* 1437-1455.

Erfurt, J. C, Foote, A., & Heirich, M. A. (1992). The cost-effectiveness of worksite wellness programs for hypertension control, weight loss, smoking cessation, and exercise. *Personnel Psychology, 45,* 5-27.

Eron, L. D., Huesman, L. R., Lefkowitz, M. M., & Walder, L. O. (1972). Does television violence cause aggression? *American Psychologist, 27,* 253-263.

Eysenck, H.J. (1952). The effects of psychotherapy: An *Qvaluation.Journal of Consulting Psychology, 16,319-32A.*

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent.* New York: Oxford University Press.

Fancher, R. E. (1990). *Pioneers of psychology* (2nd ed.). New York: W.W. Norton.

Fechner, G. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt. (Original work published 1860).

Feinberg, J. (1974). The rights of animals and unborn generations. In W. T. Blackstone (Ed.), *Philosophy and environmental crisis* (pp. 43-68). Athens: University of Georgia Press.

Fernald, D. (1984). *The Hans legacy.* Hillsdale, NJ: Erlbaum.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* Englewood Cliffs, NJ: Prentice-Hall.

Festinger, L. (1999). Reflections on cognitive dissonance: 30 years later. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 381-385). Washington, DC: American Psychological Association.

Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails.* Minneapolis: University of Minnesota Press.

Fink, A. (1995). *How to ask survey questions.* Thousand Oaks, CA: Sage.

Fischman, M. W. (2000). Informed consent. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 35-48). Washington, DC: American Psychological Association.

Fisher, R. A. (1925). *Statistical methods for research workers.* London: Oliver & Boyd.

Fisher, R. A. (1951). *The design of experiments* (6th ed.). New York: Hafner. (Original work published 1935).

Fisher, R. A., & Yates, F. (1963). *Statistical tables for biological, agricultural, and medical research* (6th ed.). Edinburgh: Oliver & Boyd.

Fisher, C. B., & Fryberg, D. (1994). Participant partners: College students weight the costs and benefits of deceptive research. *American Psychologist, 49,* 417-427.

Fisher, R. P, Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime, *journal of Applied Psychology, 74,* 722-727.

Flourens, P. (1978). Phrenology examined (C. D. Meigs, Trans.). In D. N. Robinson (Ed.), *Significant contributions to the history of psychology.* Series E.Volume II. Washington, DC: University Publications of America. (Original work published 1846).

Folkman, S. (2000). Privacy and confidentiality. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 49-57). Washington, DC: American Psychological Association.

Forrest, D. W. (1974). *Francis Galton: The life and work of a Victorian genius.* New York: Taplinger.

Fowler, F.J. Jr. (1993). *Survey research methods* (2nd ed.). Newbury Park, CA: Sage.

Foxx, R. M., & Rubinoff, A. (1979). Behavioral treatment of caffeinism: Reducing excessive coffee drinking. *Journal of Applied Behavior Analysis, 12,* 335-344.

Fraysse, J. C, & Desprels-Fraysse A. (1990). The influence of experimenter attitude on the performance of children of different cognitive ability levels. *Journal of Genetic Psychology, 151,* 169-179.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Tomlinson-Keasey, C, Martin, L. R., Wingard, D. L., & Criqui, M. H. (1995). Psychological and behavioral predictors of longevity: The aging and death of the «Termites.» *American Psychologist, 50,* 69-78.

Friedman, M., & Rosenman, R. H. (1974). *Type A behavior and your heart.* New York: Knopf.

Fruchter, B. (1954). *Introduction to factor analysis.* Princeton, NJ: DA/'an Nostrand.

Fulero, S. M., & Kirkland, J. (1992, August). *A suivey of student opinions on animal research.* Poster presented at the annual meeting of the American Psychological Association, Washington, DC.

Gallup, G. G., & Beckstead, J. W (1988). Attitudes toward animal research. *American Psychologist, 43,* 74-76.

Gallup, G. G., & Eddy, T.J. (1990). Animal facilities survey. *American Psychologist, 45,* 400-401.

Gallup, G. G., & Suarez, S. D. (1985a). Animal research versus the care and maintenance of pets: The names have been changed but the results remain the same. *American Psychologist, 40,* 968.

Gallup, G. G., & Suarez, S. D. (1985b). Alternatives to the use of animals in psychological research. *American Psychologist, 40,* 1104-1111.

Galton, F. (1872). Statistical inquiries into the efficacy of praver. *Fortnightly Review, 12,* 125-135.

Galton, F. (1883). *An inquiry into human faculty and its development.* London: MacMillan.

Gardner, G. T. (1978). Effects of federal human subjects regulations on data obtained in environmental stressor research. *Journal of Personality and Social Psychology, 36,* 628-634.

Geller, E. S. (1991). Editor's introduction: Where's the validity in social validity? *Journal of Applied Behavior Analysis, Monograph 5,* 1-6.

Gibson, E.J. (1980). Eleanor J. Gibson. In G. Lindsey (Ed.), *A history of psychology in autobiography.* Volume 7 (pp. 239-271). San Francisco: W H. Freeman.

Gibson, E. J., & Walk, R. D. (1960). The «visual cliff.» *Scientific American, 202,* 64-71.

Gilchrist, V. J., & Williams, R. L. (1999). Key informant interviews. In B. F. Crabtree & W. L. Miller (Eds.), *Doing qualitative research* (2nd ed., pp. 71-88).Thousand Oaks, CA: Sage.

Gillespie, R. (1988). The Hawthorne experiments and the politics of experimentation. In J. G. Morawski (Ed.), *The rise of experimentation in American psychology* (pp. 114-137). New Haven, CT: Yale University Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Gleitman, H., Fridlund, A. J., & Riesberg, D. (1999). *Psychology* (5th ed.). New York: W. W. Norton.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66,* 325-331.

Goodall, J. (1978). Chimp killings: Is it the man in them? *Science News, 113,* 276.

Goodwin, C.J. (1985). On the origins of Titchener's Experimentalists. *Journal of the History of the Behavioral Sciences, 21,* 383-389.

Goodwin, C.J. (1999). *A history of modem psychology.* New York: Wiley.

Green, B. F. (1992). Expose or smear? The Burt affair. *Psychological'Science, 3,* 328-331.

Greenberg, M. S. (1967). Role playing: An alternative to *deception.Journal of Personality and Social Psychology, 7,* 152-157.

Greenwald, A. G., Spangenberg, E. R., Pratkanis, A. R., & Eskenazi, J. (1991). Double-blind tests of subliminal self-help audiotapes. *Psychological Science, 2,* 119-122.

Grissett, N. I., & Norvell, N. K. (1992). Perceived social support, social skills, and quality of relationships in bulimic women. *Journal of Consulting and Clinical Psychology, 60,* 293-299.

Grose, P. (1994). *Gentleman spy: The life of Allen Dulles.* Boston: Houghton-Mifflin.

Gugerty, L.J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3,* 42-66.

Gunter, B., Berry, C, & Clifford, B. R. (1981). Proactive interference effects with television news items: Further evidence. *Journal of Experimental Psychology: Human Learning and Memory,* 7,480-487.

Gwaltney-Gibbs, P. A. (1986). The institutionalization of premarital cohabitation: Estimates from marriage license applications, 1970 and 1980.*Journal of Marriage and the Family, 48,* 423-434.

Hall, G. S. (1893). *The contents of children's minds on entering school.* New York: Kellogg.

Hall, J. A., & Veccia, E. M. (1990). More «touching» observations: New insights on men, women, and interpersonal touch. *Journal of Personality and Social Psychology, 59,* 1155-1162.

Harris, B. (1979). Whatever happened to little Albert? *American Psychologist,34,* \5\~\60.

Hartman, D. P., & Hall, R. V. (1976).The changing criterion design. *Journal of Applied Behavior Analysis, 9.* 527-532.

Henle, M., & Hubbell, M. B. (1938). «Egocentricity» in adult conversation. *Journal of Social Psychology, 9,* 227-234.

Hilgard, E. R. (Ed.). (1978). *American psychology in historical perspective.* Washington, DC: American Psychological Association.

Hilgard, E. R. (1987). *Psychology in America: A historical survey.* San Diego, CA: Harcourt Brace Jovanovich.

Hilgartner, S. (1990). Research fraud, misconduct, and the IRB. *IRB:A Review of Human Subjects Research, 12,* 1-4.

Hite, S. (1987). *Women and love.* New York: Knopf.

Hobbs, N. (1948). The development of a code of ethics for psychology. *American Psychologist, 3,* 80-84.

Holahan, C. K., Sears, R. R..& Cronbach, L.J. (1987). *The gifted group in midlife.* Stanford, CA: Stanford University Press.

Holden, C (1987, March 27). NIMH finds a case of «serious misconduct.» *Science, 235,* 1566-1577.

Holmes, D. S. (1976a). Debriefing after psychological experiments. I. Effectiveness of postdeception dehoaxing. *American Psychologist, 31,* 858-867.

Holmes, D. S. (1976b). Debriefing after psychological experiments. II. Effectiveness of postexperimental desensitizing. *American Psychologist, 31,* 868-875.

Holmes, D. S„ Mcgilley, B. M., & Houston, B. K. (1984). Task-related arousal and Type A and Type B persons: Levels of challenge and response specificity. *Journal of Personality and Social Psychology, 46,* 1322-1327.

Horn, J. L. (1990, October). Psychology can help kids get a Head Start. *APA Monitor, 22,3.*

Hothersall, D. (1990). *History of psychology* (2nd ed.). New York: McGraw-Hill.

Howell, D. C (1997). *Statistical methods for psychology* (4th ed.). Belmont, C A: Duxbury Press.

Hubel, D. H. (1988). *Eye, brain, and vision.* New York: Scientific American Library.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate *zoxtex. Journal of Physiology, 148,* 574-591.

Huff, D. (1954). *How to lie with statistics.* New York: W.W. Norton.

Hull, D. B. (1996). Animal use in undergraduate psychology programs. *Teaching of Psychology, 23,* 171-174.

Hume, K. M., & Grossman, J. (1992). Musical reinforcement of practice behaviors among competitive swimmers. *Journal of Applied Behavior Analysis, 25,* 665-670.

Hunt, E., & Love, T. (1972). How good can memory be? In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 237-260). Washington, DC:V. H. Winston.

Jenkins, J. G., & Dallenbach, K. M. (1924). Minor studies from the psychological laboratory of Cornell University: Oblivescence during sleep and waking. *American Journal of Psychology, 35,* 605-612.

Joncich, G. (1968). *The sane positivist: A biography of Edward L. Thomdike.* Middletown, CT: Wesleyan University Press.

Jones, J. H. (1981). *Bad blood: The Tuskegee syphilis experiment.* New York: Free Press.

Jones, M. C. (1924). A laboratory study of fear: The case of Peter. *Pedagogical Seminary, 31,* 308-315.

Jordan, N. C, Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28,* 644-653.

Junginger, J., & Head, S. (1991). Time series analyses of obsessional behavior and mood during self-imposed delay and responsive prevention. *Behavior Research and Therapy, 29,* 521-530.

Kaufman, A. S., & Kaufman, N. L. (1983). *KABC: Kaufman Assessment Battery for Children. Interpretive manual.* Circle Pines, MN: American Guidance Service.

Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research.* Baltimore: University Park Press.

Keltner, D., Ellsworth, P. C, & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology, 64,* 740-752.

Kendall, M. G. (1970). Ronald Aylmer Fisher, 1890-1962. In E. S. Pearson & M. G. Kendall (Eds.), *Studies in the history of statistics and probability.* London: Charles Griffin.

Kent, D. (1994). Interview with APS president-elect Richard F. Thompson. *APS Observer, 7, 4,* 10.

Key, W. B. (1973). *Subliminal seduction.* Englewood Cliffs, NJ: Signet.

Kidd, V. (1971). A study of the images produced through the use of the male pronoun as the generic. *Moments in Contemporary Rhetoric and Communication, 1,* 25-30.

Kim, K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object *motion. Journal of Experimental Psychology: Human Perception and Performance, 18,* 385-393.

Kimmel, A. J. (1996). *Ethical issues in behavioral research: A survey.* Cambridge, MA: Blackwell.

Kinsey, A. C, Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male.* Philadelphia: W. B. Saunders.

Kinsey, A. C, Pomeroy, W. B., Martin, C E., & Gebhard, P. H. (1953). *Sexual behavior in the human female.* Philadelphia: W. B. Saunders.

Kirk, R. E. (1968). *Experimental design: Procedures for the behavioral sciences.* Belmont, CA: Brooks/Cole.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism *mjapanjournal ofPersonality and Social Psychology, 72,* 1245-1267.

Knepper, W., Obrzut, J. E., & Copeland, E. P. (1983). Emotional and social problem-solving thinking in gifted and average elementary school *children.Joumal of Genetic Psychology, 142,* 25-30.

Kohlberg, L. (1963). Development of children's orientation towards a moral order. *Vita Humana, 6,* 11-36.

Kohn, A. (1986). *False prophets: Fraud and error in science and medicine.* Oxford: Basil Blackwell.

Kolata, G. B. (1986).What does it mean to be random? *Science, 231,* 1068-1070.

Korn, J. H. (1988). Students' roles, rights, and responsibilities as research participants. *Teaching of Psychology, 15,* 74-78.

Korn, J. H., Davis, R., & Davis, S. F. (1991). Historians' and chairpersons'judgments of eminence among psychologists. *American Psychologist, 46,* 789-792.

Kosslyn, S. M., Ball, T. M.,& Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 47-60.

Krantz, D. L. (1969). The Baldwin-Titchener controversy: A case study in the functioning and malfunctioning of schools. In D. L. Krantz (Ed.), *Schools of psychology: A symposium* (pp. 1-19). New York: Appleton-Century-Crofts.

Krupat, E. (1975). Conversation with John Darley. In E. Krupat (Ed.), *Psychology is social: Readings and conversations in social psychology.* Glen view, IL: Scott, Foresman.

Kruta, V. (1972). Marie-Jean-Pierre Flourens. In C. C. Gillespie (Ed.), *Dictionary of scientific biography* (Vol.V). New York: Scribner's.

Kuhn, T. S. (1970).The function of dogma in scientific research. In B. A. Brody (Ed.), *Readings in the philosophy of science* (pp. 356-373). Englewood Cliffs, NJ: Prentice-Hall.

Kushner, M. (1970). Faradic aversive controls in clinical practice. In C Neuringer & J. L. Michael (Eds.), *Behavior modification in clinical practice.* New York: Appleton-Century-Crofts.

Landrum, E. R., & Chastain, G. (1999). Subject pool policies in undergraduate-only departments: Results from a nationwide survey. In G. Chastain & E. R. Landrum (Eds.), *Protecting human subjects: Department subject pools and Institutional Review Boards* (pp. 25-42). Washington, DC: American Psychological Association.

Landy, F.J. (1992). Hugo Mbnsterberg: Victim or *visionary? Journal of Applied Psychology, 77,* 787-802.

Langer, E. J., & Rodin, J. (1976).The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology, 34,* 191-198.

Lau, R. R., & Russell, D. (1980). Attributions in the sports pages. *Journal of 'Personality and Social Psychology, 39,* 29-38.

Lavkakas, P.J. (1998). Methods for sampling and interviewing in telephone surveys. In L. Bickman & D.J. Rog (Eds.), *Handbook of applied social research methods* (pp. 429-472).Thousand Oaks, CA: Sage.

Leak, G. K. (1981). Student perception of coercion and value from participation in psychological research. *Teaching of Psychology, 8,* 147-149.

Lee, D. N., & Aronson, E. (1974).Visual proprio-ceptive control of standing in human infants. *Perception andPsychophysics, 15,* 529-532.

Lefrancois, J. R., & Metzger, B. (1993). Low-response-rate conditioning history and fixed-interval responding in *Tuts. Journal of the Experimental Analysis of Behavior, 59,* 543-549.

Liberman, R. P., Davis, J., Moon, W., & Moore, J. (1973). Research design for analyzing drug-environment-behavior *interactions Journal of Newous and Mental Disease, 156,* 432-439.

Licht, M. M. (1995). Multiple regression and correlation. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics {pp.* 19-64). Washington, DC: American Psychological Association.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F. (1986). Ten years in the life of an expert witness. *Law and Human Behavior, 10,241-263.*

Loftus, E. F, & Hoffman, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118,* 100-104.

Loftus, E. F, & Ketcham, K. (1991). *Witnessforthe defense: The accused, the eyewitness, and the expert who puts memory on trial.* New York: St. Martins Press.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585-589.

Lorenz, K. (1966). *On aggression.* New York: Har-court Brace Jovanovich.

Lotufo, P. A., Chae, C U., Ajani, U. A., Hennekens, C. H., & MansonJ. E. (1999). Male pattern baldness and coronary heart disease: The physician's health study. *Archives of Internal Medicine, 160,* 165-171.

Ludwig, T. D., & Geller, E. S. (1997). Assigned versus participative goal setting and response generalization: Managing injury control among professional pizza deliverers.*Journal of'Applied Psychology, 82,* 253-261.

Luria, A. R. (1968). *The mind of a mnemonist.* New York: Basic Books.

Macleod, C. M. (1992). The Stroop task: The «gold standard» of attentional measures. *Journal of Experimental Psychology: General, 121,* 12-14.

Mangione, T. W. (1998). Mail surveys. In L. Bickman & D.J. Rog (Eds.), *Handbook of applied social research methods* (pp. 399-427).Thousand Oaks, C A: Sage.

Manis, M. (1971). *An introduction to cognitive psychology.* Belmont, C A: Brooks/Cole.

Marcel, A., J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197-237.

Marean, G. C, Werner, L. A., & Kuhl, P. K. (1992). Vowel categorization by very young infants. *Developmental Psychology, 28,*396-405.

McClelland, D. C (1961). *The achieving society.* Princeton, NJ: Van Nostrand.

McClelland, D. C, Atkinson, J. W., Clarke, R. A., & Lowell, E. L. (1953). *The achievement motive.* New York: Appleton-Century-Crofts.

McReynolds, P. (1987). Lightner Witmer: Little-known founder of clinical psychology. *American Psychologist, 42,* 849-858.

Merikle, P. M., & Skanes, H. E. (1992). Subliminal self-help audiotapes: A search for placebo effects. *Journal o/ Applied Psychology, 77,* 772-776.

Middlemist, R. D., Knowles, E. W., & Matter, C F. (1976). Personal space invasions in the lavatory: Suggestive evidence for arousal. *Journal of Personality and Social Psychology, 33,* 541-546.

Miles, W. R. (1928). Studies in physical exertion I: A multiple chronograph for measuring groups of men. *American Physical Education Review, 33,* 379-387.

Miles, W. R. (1930). On the history of research with rats and mazes: A collection of notes. *Journal of General Psychology, 3,* 324-337.

Miles, W. R. (1931). Studies in physical exertion II: Individual and group reaction time in football charging. *Research Quarterly, 2(3),* 5-13.

Miles, W. R. (1933). Age and human ability. *Psychological Review, 40,* 99-123.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371-378.

Milgram, S. (1974). *Obedience to authority: An experimental view.* New York: Harper & Row.

Miller, A, G. (1972). Role playing: An alternative to deception? *American Psychologist, 27,* 623-636.

Miller, N. (1985). The value of behavioral research on animals. *American Psychologist, 40,* 423-440.

Miller, W. R., & Dipilato, M. (1983). Treatment of nightmares via relaxation and desensitization: A controlled evaluation./ownW *of Consulting and Clinical Psychology, 51,* 870-877.

Minton, H. L. (1987). Lewis M. Terman and mental testing: In search of the democratic ideal. In M. M. Sokal (Ed.), *Psychological testing and American society.* 1890-1930 (pp. 95-112). New Brunswick, NJ: Rutgers University Press.

Minton, H. L. (1988). Charting life history: Lewis M. Terman's study of the gifted. In J. G. Morawski (Ed.), *The rise of experimentation in American psychology* (pp. 138—162). New Haven, CT: Yale University Press.

Mischel, W. (1981). *Introduction to personality* (3rd ed.). New York: Holt, Rinehart, & Winston.

Morgan, C L. (1903). *Introduction to comparative psychology.* London: Walter Scott.

Morgan, F. W (1990). Judicial standards for survey research: An update and guidelines. *Journal of Marketing, 54,* 59-70.

Moses, S. (1991, July). Animal research issues affect students. *APA Monitor, 22,* 47-48.

Mowrer, O. H., & Mowrer, W. M. (1938). Enuresis — a method for its study and treatment. *American Journal of Orthopsychiatry, 8,* 436-459.

Murray, H. A. (1943). *Thematic apperception test.* Cambridge, MA: Harvard University Press.

Myers, D. G. (1990). *Social psychology* (3rd ed). New York: McGraw-Hill.

Myers, D. G. (1992). *Psychology* (3rd ed.). New York: Worth.

Neale, J. M., & Liebert, R. M. (1973). *Science and behavior: An introduction to methods of research.* Englewood Cliffs, NJ: Prentice-Hall.

Neisser, U. (1963). Decision-time without reaction-time. *American Journal of Psychology, 76,* 376-385.

Neisser, U. (1976). *Cognition and reality.* San Francisco: W. H. Freeman.

Neisser, U. (1981). John Dean's memory: A case study. *Cognition, 9,* 1-22.

NorcrossJ. C, HanychJ. M., & Terranova, R. D. (1996). Graduate study in psychology: 1992-1993. *American Psychologist, 51,* 631-643.

O'Donnell, J. M. (1985). *The origins of behaviorism: American psychology, 1870-1920.* New York: New York University Press.

Orne, M. T. (1962). On the social psychology of the psychology experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17,* 776-783.

Orne, M. T., & Scheibe, K. E. (1964). The contribution of nondeprivation factors in the production of sensory deprivation *efiects.Journal ofAbnormal and Social Psychology, 68,3-12.*

Orwin, R. G. (1997). Twenty-one years old and counting: The interrupted time series comes of age. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation forthe 2'1st century* (pp. 443-465). Thousand Oaks, CA: Sage.

Osier, S. P., & Trautman, G. E. (1961). Concept attainment II: Effect of stimulus complexity upon concept attainment at two levels of intelligence. *Journal of Experimental Psychology, 62,* 9-13.

Overmier, J. B., & Leaf, R. C (1965). Effects of discriminative Pavlovian fear conditioning upon previously or subsequently acquired avoidance *responding. Journal of Comparative and Physiological Psychology, 60,* 213-218.

Parson, H. M. (1974). What happened at Hawthorne? *Science, 183,* 922-932.

Patten, M. L. (1998). *Questionnaire research: A practical guide.* Los Angeles: Pryczak.

Patterson, F. G„ & Linden, E. (1981). *The education ofKoko.* New York: Holt, Rinehart, & Winston.

Pearson, E. S., & Hartley, H. O. (1966). *Biometrika tables for statisticians* (Vol. 1, 3rd ed.). London: Cambridge University Press.

Peck, F. S. (1978). *The road less traveled.* New York: Simon & Schuster.

Peterson, C, Maier, S. F, & Seligman, M. E. P. (1993). *Learned helplessness: A theory for the age of personal control.* New York: Oxford University Press.

Peterson, L., Ridley-Johnson, R., & Carter, C. (1984). The supersuit: An example of structured naturalistic observation of children's altruism. *Journal of General Psychology, 110,235-241.*

Pious, S. (1996a). Attitudes toward the use of animals in psychological research and education: Results from a national survey of psychologists. *American Psychologist, 51,* 1167-1180.

Pious, S. (1996b). Attitudes toward the use of animals in psychological research and education: Results from a national survey of psychology majors. *Psychological Science, 7,* 352-358.

Popper, K. R. (1959). *The logic of scientific discovery.* New York: Basic Books.

Porsolt, R. D., Lepichon, M., & Jalfre, M. (1977). Depression: A new animal model sensitive to antidepressant treatments. *Nature, 266,* 730-732.

Posavac, E.J., & Carey, R. G. (1997). *Program evaluation: Methods and case studies* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Poulton, E. C (1982). Influential companions: Effects of one strategy on another in the within-subjects designs of cognitive psychology. *Psychological Bulletin, 91,*673-690.

Pratkanis, A. R., Eskenazi, J., & Greenwald, A. G. (1994). What you expect is what you believe (but not necessarily what you get): A test of the effectiveness of subliminal self-help audiotapes. *Basic and Applied Social Psychology, 15,* 251-276.

Randi, J. (1982). *Film flam! Psychics, ESP, unicorns, and other delusions.* Buffalo, NY: Prometheus Books.

Rauscher, F. W., Shaw, G. L., & Key, K. N. (1993). Music and spatial task performance. *Nature, 365,* 611.

Reeve, L., Reeve, K. F., Brown, A. K., Brown, J. L., & Poulson, C L. (1992). Effects of delayed reinforcement on infant vocalization *rate. Journal of the Experimental Analysis of Behavior, 58,* 1-8.

Renne, C M., & Creer, T. L. (1976). Training children with asthma to use inhalation therapy equipment.yoMma/ *of Applied Behavior Analysis, 9,* 1-11.

Reynolds, G. S. (1968). *A primer ofoperant conditioning.* Glenview, IL: Scott, Foresman.

Reynolds, R. I. (1992). Recognition of expertise in chess players. *American Journal of Psychology, 105,* 409-415.

Riskind, J. H., & Maddux, J. E. (1993). Loomingness, helplessness, and fearfulness: An integration of harm-looming and self-efficacy models of fear. *Journal of Social and Clinical Psychology, 12,* 73-89.

Rockefeller, J. D. IV. (1994). *Is military research hazardous to veterans' health? Lessons spanning half a century: A repo?t examining biological experimentation on U.S. military* [on-line].www.trufax.org/trans/roc00.html

Rodin, J., & Langer, E.J. (1977). Long-term effects of a control-relevant intervention with the institutionalized aged. *Journal of Personality and Social Psychology, 35,* 897-902.

Rogelberg, S. G., & Luong, A. (1998). Nonre-sponse to mailed surveys: A review and guide. *Current Directions in Psychological Science, 7,* 60-65.

Rogosa, D. (1980). A critique of cross-legged correlation. *Psychological Bulletin, 88,*245-258.

Rohles, F. H., Jr. (1992). Orbital bar pressing: A historical note on Skinner and the chimpanzees in space. *American Psychologist, 47,* 1531-1533.

Romanes, G.J. (1886). *Animal intelligence.* New York: D. Appleton.

Rosenberg, M.J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 280-349). New York: Academic Press.

Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.

Rosenthal, R. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science, 5,* 127-133.

Rosenthal, R., & Fode, K. L. (1963a). Three experiments in experimenter bias. *Psychological Reports, 12,* 491-511.

Rosenthal, R., & Fode, K. L. (1963b). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science, 8,* 183-189.

rosenthal, R., & rosnow, R. L. (1991). Essentials of behavioral research: Methods and data analysis (2nd ed.). New York: McGraw-Hill.

Rosnow, R. L., Goodstadt, B. E., Suls, J. M., & Gitter, A. G. (1973). More on the social psychology of the experiment: When compliance turns to self-defense, *Journal of Personality and Social Psychology, 27,* 337-343.

Rucci, A. J., & Tvveney, R. D. (1980). Analysis of variance and the «second discipline» of scientific psychology: A historical account. *Psychological Bulletin, 87,* 166-184.

Samelson, F. (1992). Rescuing the reputation of Sir Cyril [Burt]. *Journal of the History of the Behavioral Sciences, 28,* 221-233.

Sanford, E. C. (1910, August 8). *Letter to E. B. Titchener.* Ithaca, NY: Titchener Papers, Cornell University.

Sanford, E. C. (1914). Psychic research in the animal field: Der Kluge Hans and the Elberfeld horses. *American Journal of Psychology, 25,* 3-31.

Schachter, S. (1959). *The psychology of affiliation.* Stanford, CA: Stanford University Press.

Schaie, K. W. (1983). The Seattle longitudinal study: A 21 year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development.* New York: Guilford Press.

Schaie, K. W. (1988). Ageism in psychological research. *American Psychologist, 43,*179-183.

Schneider, W., & Bjorklund, D. F. (1992). Expertise, aptitude, and strategic remembering. *Child Development, 63,* 461-473.

Schneirla, T. C (1929). Learning and orientation in ants. *Comparative Psychology Monographs, 6* (No. 4).

Schoeneman, T. J., & Rubanowitz, D. E. (1985). Attributions in the advice columns: Actors and observers, causes and reasons. *Personality and Social Psychology Bulletin,* //,315-325.

Schrader, w. B. (1971). The predictive validity of the College Board Admissions tests. In W. H. Angoff (Ed.), *The College Board Admission Testing Program.* New York: College Entrance Examination Board.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and content.* Thousand Oaks, C A: Sage.

Scott-Jones, D. (2000). Recruitment of research participants. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 27-34). Washington, DC: American Psychological Association.

Scripture, E. W. (1895). *Thinking, feeling, doing.* Meadville, PA: Chautauqua-Century Press.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515-530.

Sechrest, L., & Figueredo, A. J. (1993). Program evaluation. In L. W. Porter & M. R. Rosenzweig(Eds.), *Annual review of psychology. Volume 44 (pp.* 645-674). Palo Alto, CA: Annual Reviews.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: W. H. Freeman.

Seligman, M. E. P., Castellon, C, Cacciola, J., Schulman, P., Luborsky, L., Ollove, M., & Downing, R. (1988). Explanatory style change during cognitive therapy for unipolar depression. *Journal of Abnormal Psychology, 97,*13-18.

Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology, 74,* 1-9.

Seligman, M. E. P., & Schulman, P. (1986). Explanatory style as a predictor of productivity and quitting among life insurance agents. *Journal of Personality and Social Psychology, 50,* 832-838.

Sheldon, W. H. (1940). *The varieties of human physique: An introduction to constitutional psychology.* New York: Harper & Row.

Sheldon, W. H. (1942). *The varieties oftemperament: A psychology of constitutional differences.* New York: Harper & Row.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,701-703.*

Sherrod, D. R., Hage, J. N., Halpern, P. L., & Moore, B. S. (1977). Effects of personal causation and perceived control on responses to an environment: The more control, the better. *Journal of Personality and Social Psychology, 13,* 14.

Shute, V.J. (1994). Learners and instruction: What's good for the goose may not be good for the gander. *Psychological Science Agenda, 7(3),* 8-9, 16.

Sidman, M. (1960). *Tactics of scientific research.* New York: Basic Books.

Sieber, J. E. (1994). Will the new code help researchers to be more ethical? *American Psychologist, 25,* 369-375.

Sieber, J. E. (1998). Planning ethically responsible research. In L. Bickman & D.J. Rog (Eds.), *Handbook of applied social research methods* (pp. 127-156). Thousand Oaks, CA: Sage.

Sieber, J. E., & Saks, M.J. (1989). A census of subject pool characteristics. *American Psychologist, 44,* 1053-1061.

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime *on juridic judgment. Journal of Personality and Social Psychology, 31,* 410-414.

Silverman, I. (1975). Nonreactive methods and the law. *American Psychologist, 30,*764-769.

Singer, P. (1975). *Animal liberation.* New York: Avon.

Skinner, B. F. (1953). *Science and human behavior.* New York: Free Press.

Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 12,* 221-233.

Skinner, B. F. (1966). Operant behavior. In W. K. Honig(Ed.), *Operant behavior: Areas of research and application* (pp. 12-32). New York: Appleton-Century-Crofts.

Skinner, B. F. (1969). *Contingencies of reinforcement.* Englewood Cliffs, NJ: Prentice-Hall.

Skinner, B. F. (1976). *Walden Two.* New York: Macmillan. (Original work published 1948).

Skinner, B. F. (1979). *The shaping of a behaviorist.* New York: New York University Press.

Skinner, B. F. (1984). *A matter of consequences.* New York: New York University Press.

Small, W. S. (1900). An experimental study of the mental processes of the rat. *American Journal of Psychology, 11,* 80-100.

Smith, L. D. (1992). On prediction and control: B. F. Skinner and the technological ideal of science. *American Psychologist, 47,* 216-223.

Smith, S., & Sechrest, L. (1991). Treatment of aptitude X treatment interactions.yowraa/ *of Consulting and Clinical Psychology, 59,* 233-244.

Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research: The important role *of dehneiing.Journal of Personality and Social Psychology, 44,* 1075-1082.

Smoll, F. L., Smith, R. E., Barnett, N. P., & Everett, J. J. (1993). Enhancement of children's self-esteem through social support training for youth sport coaches. *Journal of Applied Psychology, 78,* 602-610.

Society for Research in Child Development. (1996). *Ethical standards for research with children.* SRCD Directory of Members, 337-339.

Sokal, M. M. (Ed.) (1981). *An education in psychology:James McKeen Cattell's journal and letters from Germany and England, 1880-1888.* Cambridge, MA: MIT Press.

Sokal, M. M. (1987). James McKeen Cattell and mental anthropometry: Nineteenth-century science and reform and the origins of psychological testing. In M. M. Sokal (Ed.), *Psychological testing and American society, 1890-1930* (pp. 21-45). New Brunswick, NJ: Rutgers University Press.

Sokal, M. M. (1992). Origins and early years of the American Psychological Association, 1890-1906. *American Psychologist, 47,* 111-122.

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin, 46,* 137-150.

Spatz, C (1997). *Basic statistics: Tales of distributions* (6th ed.). Pacific Grove, CA: Brooks/Cole.

Spelke, E. S. (1985). Preferential looking methods as tools for the study of cognition in infancy. In G. Gottlieb & N. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life* (pp. 323-363). Norwood, NJ: Ablex.

Sprinthall, R. C. (2000). *Basic statistical analysis* (6th ed.). Boston: Allyn & Bacon.

Spurzheim, J. G. (1978). Outlines of phrenology. In D. N. Robinson (Ed.), *Significant contributions to the history of psychology.* Series E. Volume II. Washington, DC: University Publications of America. (Original work published 1832).

Steele, K. M., Ball, T. N., & Runk, R. (1997). Listening to Mozart does not enhance backwards digit span performance. *Perceptual and Motor Skills, 84,* 1179-1184.

Sternberg, R. J., & Grigorenko, E. L. (1999). A smelly 113° in the shade (Or, why we do field research). *APS Observer, 12(8),* 10-11, 20-21.

Sternberg, R. J., Powell, C., McGrane, P., & Grantham-McGregor, S. (1997). Effects of aparasitic infection on cognitive functioning. *Journal of Experimental Psychology: Applied, 3,* 67-76.

Stolzenberg, L., & D'Alessio, S.J. (1997). «Three strikes and you're out»: The impact of California's new mandatory sentencing law on serious crime rates. *Crime and Delinquency, 43,* 457-470.

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 121,* 15-23. (Original work published 1935).

Sullivan, D. S., & Deiker, T. E. (1973). Subject-experimenter perceptions of ethical issues in human research. *American Psychologist, 28,* 587-591.

Tatham, T. A., Wanchisen, B. A., & Hineline, P. N. (1993). Effects of fixed and variable ratios on human behavioral *variability. Journal of the Experimental Analysis of Behavior, 59,*349-359.

Taylor, D. W., Garner, W. R., & Hunt, H. F. (1959). Education for research in psychology. *American Psychologist, 14,* 167-179.

Taylor, S. J., & Bogdan, R. (1998). *Introduction to qualitative methods: A guide and resource* (3rd ed.). New York: Wiley.

Terman, L. M. (1925). *Genetic studies of genius, Vol. 1. Mental and physical traits of a thousand gifted children.* Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1947). *Genetic studies of genius, Vol. 4. The gifted child-grows up: Twenty-five years' follow-up of a superior group.* Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius, Vol. 5. The gifted group at mid-life: Thirty-five years' follow-up of the superior child.* Stanford, CA: Stanford University Press.

Terrace, H. (1963). Discrimination learning with and without «errors». *Journal of the Experimental Analysis of Behavior, 6,* 1-27.

Thomas, E. (1995). *The very best men.* New York: Simon & Schuster.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monographs, 2* (No. 8).

Thorndike, E. L. (2000). *Animal intelligence: Experimental studies.* New Brunswick, NJ: Transaction Publishers. (Original work published 1911).

Thurstone, L. L. (1938). *Primary mental abilities. Psychometric Monographs, No. 1.* Chicago: University of Chicago Press.

Titchener, E. B. (1916). *A textbook of psychology.* New York: MacMillan. (Original work published 1909).

Todd, J. T., & Morris, E. K. (1992). Case histories in the great power of steady misrepresentation. *American Psychologist, 47,* 1441-1453.

Tolman, E. C (1959). Principles of purposive behavior. In S. Koch (Ed.), *Psychology: A study of a science: Volume 2. General systematic formulations, learning, and special processes* (pp. 92-157). New York: McGraw-Hill.

Tolman, E. C, Tryon, R. C, & Jeffries, L. A. (1929). A self-recording maze with an automatic delivery table. *University of California Publications in Psychology, 4,* 99-112.

Tomas, V. (Ed.). (1957). *Charles S. Peirce: Essays in the philosophy of science.* New York: Liberal Arts Press.

Tomlinson-Keasey, C (1990). The working lives of Terman's gifted women. In H. Y. Grossman & N. L. Chester, (Eds.), *The experience and meaning of work in women's lives* (pp. 213-239). Hillsdale, NJ: Lawrence Erlbaum.

Trabasso, T. (1963). Stimulus emphasis and all-or-none learning in concept identification. *Journal of Experimental Psychology, 65,* 398-406.

Triandis, H. C (1995). *Individualism and collectivism.* Boulder, CO: Westview Press.

Tryon, R. C (1929). The genetics of learning ability in rats: Preliminary report. *University of California Publications in Psychology, 4,* 71-89.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Tulving, E. (1966). Subjective organization and the effects of repetition in multi-trial free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 195-197.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic forjudging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Tweney, R. D. (1987). Programmatic research in experimental psychology: E. B. Titchener's laboratory investigations, 1891-1927. In M. G. Asch & W. R. Woodward (Eds.), *Psychology in twentieth-century thought and society* (pp. 34-57). New York: Cambridge University Press.

Ulrich, R. S. (1984).View through a window may influence recovery from surgery. *Science, 224,* 420-421.

van Kammen, W. B., & Stouthamer-Loeber, M. (1998). Practice aspects of interview data collection and data management. In L. Bickman & D.J. Rog (Eds.), *Handbook of applied social research methods* (pp. 375-397). Thousand Oaks, CA: Sage.

Wagaman, J. R., Miltenberger, R. G., & Arndorfer, R. E. (1993). Analysis of a simplified treatment for stuttering in *children.Journal of Applied Behavior Analysis, 26,*53-61.

Wagner, J. A., Ill, Rubin, P. A., & Callahan, T.J. (1988). Incentive payment and nonmanagerial productivity: An interrupted time series analysis of magnitude and trend. *Organizational Behavior and Human Decision Processes, 42,* 47-74.

Walker, A.J. (1996). Couples watching television: Gender, power, and the remote control. *Journal of Marriage and the Family, 58,* 813-823.

Wason, P. C, & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* Cambridge, MA: Harvard University Press.

Watson, J. B. (1907). Kinesthetic and organic sensations: Their role in the reactions of the white rat to the maze. *Psychological Review Monograph Supplements, 8* (No. 33).

Watson, J. B. (1924). *Behaviorism.* New York: W. W. Norton.

Watson, J. B. (1928). *Psychological care of infant and child.* New York: W.W. Norton.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1-14.

Webb, E. J., Campbell, D.T., Schwartz, R. D., Sechrest, L., & Grove, J. B. (1981). *Non-reactive measures in the social sciences* (2nd ed.). Boston: Houghton Mifflin.

Webster's word histories. (1989). Springfield, MA: Merriam-Webster.

Weiss, J. M. (1968). Effects of coping response on stress. *Journal of Comparative and Physiological Psychology, 65,* 251-260.

Weiss, J. M. (1977). Psychological and behavioral influences on gastrointestinal lesions in animal models. In J. D. Maser & M. E. P. Seligman (Eds.), *Psychopathology: Experimental models.* San Francisco: W. H. Freeman.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 2,* 440-445.

Winston, R. S. (1990). Robert Sessions Wood-worth and the «Columbia Bible»: How the psychological experiment was redefined. *Americanjoumal of Psychology, 103,*391-401.

Witt, L. A., &Nye, L. G. (1992). Gender and the relationship between perceived fairness of pay or promotion and job satisfaction. *Journal of Applied Psychology, 77,*910-917.

Wolf, M. M. (1978). Social validity: The case for subjective measurement, or how behavior analysis is finding its *heart. Journal of Applied Behavior Analysis, 11,* 203-214.

Wollen, K. A., Weber, A., & Lowry, D. H. (1972). Bizarreness versus interaction of images as determinants of learning. *Cognitive Psychology, 3,* 518-523.

Wood, J. M., & Bootzin, R. R. (1990). The prevalence of nightmares and their independence from anxiety. *Journal of Abnormal Psychology, 99,* 64-68.

Wood, J. M., Bootzin, R. R., Rosenhan, D., Nolen-Hoeksema, S., & Jourden, F. (1992). Effects of the 1989 San Francisco earthquake and content of nightmares. *Journal of Abnormal Psychology, 107,*219-224.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Henry Holt.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology, 10,* 109-120.

Wundt, W. (1904). *Principles of physiological psychology* (5th ed.) (E. B. Titchener, Trans.). New York: MacMillan. (Original work published 1874).

Yarmen, A. D., & Bull, M. P. (1978). Where were you when President Kennedy was assassinated? *Bulletin of the Psychonomic Society, 11,* 133-135.

Yeaton, W. H., & Sechrest, L. (1986). Use and misuse of no-difference findings in eliminating threats to validity. *Evaluation Review, 10,* 836-852.

Zeigarnik, B. (1967). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 300-314). London: Routledge and Kegan Paul. (Original work published 1927).

Алфавитный указатель

Л

М

Н

О

П

Р

С

Т

У

Ф

Х

Ч

Ш

Э

Джеймс Гудвин

Исследование в психологии: методы и планирование

3-е издание

Перевела с английского Л. Сиренко

Главный редактор	*Е. Строганова*
Заведующий редакцией	*Л. Винокуров*
Руководитель проекта	*Н. Римицан*
Выпускающий редактор	*А. Борин*
Научный редактор	*А. Худяков*
Литературный редактор	*С. Овечкин*
Художественный редактор	*Р. Яцко*
Корректоры	*М. Одинакова, Н. Соломина*
Верстка	*Т. Гусева*

Лицензия ИД №05784 от 07.09.01.

Подписано в печать 24.09.03. Формат 70Х100/16. Усл. п. л. 45.15. Тираж 3500 экз. Заказ № 727.

ООО «Питер Принт». 196105. Санкт-Петербург, ул. Благодатная, д. 67в.

Налоговая льгота - общероссийский классификатор продукции ОК 005-93, том 2; 953005 - литература учебная.

Отпечатано с готовых диапозитивов в ФГУП «Печатный двор» им. А. М. Горького
Министерства РФ по делам печати, телерадиовещания и средств массовых коммуникаций.
197110. Санкт-Петербург, Чкаловский пр., 15.

В данном учебнике, пользующемся большой популярностью в западной научной и студенческой среде, представлены материалы стандартного курса «Экспериментальная психология» для психологических факультетов высших учебных заведений. Умение надлежащим образом спланировать и провести эксперимент является одним из основных умений психолога, вне зависимости от его специализации. Автор книги подробно рассматривает методологию психологического исследования, организацию эксперимента, контроль переменных, проблемы интерпретации и представления данных. Четкая структура, хорошая подача материала, большое количество примеров реальных экспериментов, занимательный стиль изложения — вот те качества, которые характеризуют этот учебник и позволяют надеяться, что он окажет неоценимую помощь студентам в усвоении данной дисциплины.

Книга предназначена для студентов и преподавателей психологических факультетов, а также для всех интересующихся методами психологического исследования.

«Гудвин делает исследование интересным и даже веселым занятием (трудная задача для меня и других преподавателей)... Я обычно прошу студентов оценить качество учебника по 10-балльной шкале, и средняя оценка книги Гудвина — 9 баллов.»

Джейсон Л. Хикс, Университет Луизианы

ISBN 5-94723-290-1

